

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.

**THIS PAGE BLANK (USPTO)**

R70141	y148g11.s1
A1918219	tn08f10.x
BEL10202	UI-R-CA0-
BF401780	UI-R-CA0-
BI403166	MI-P-CPI-
BEL14223	UI-R-CA0-
BE416269	UI-R-CA0-
BII19093	AK023A09P
BF404955	UI-R-CA1-
AZ488750	IM0319M09
BII138537	BH138537
BII195829	C0317E11-
A1741034	wk25f108.x
BF402467	UI-R-CA0-
BE241758	TCAAPE2E02
BM193296	TCCAP1E89
BM146800	TCCAP1E73
BB099875	BB099875
BG243742	6023357051
BH305262	CH230-4B1
BH339028	CH230-89F
AL218285	Tetraodon
BM388754	UI-R-D20-
BF151027	uy91d04.x
AA544572	v167b12.r
BJ064111	BJ064111
BB481139	BB481139
AI3597000	v167b12.x

  

polylinker.	Site-1:	Not I;	Site-2:	Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for				
germinal center B cells by flow sorting (CD20+ IgD-),				
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman				
(NCI) and Dr. Gerald Matti (CSBR). cDNA synthesis was				
primed with a Not I - oligo(dt) primer				
[5'-GTTCACATCCTAGAGCGGCAGGCCCTCATTTTTTTTTTTTTTT-3'				
]. Double-stranded cDNA was ligated to Eco RI adaptors				
(Pharmacia), digested with Not I and cloned into the Not I				
and Eco RI sites of the modified pTV73 vector. Library				
went through one round of normalization, and was				
constructed by Bento Soares and M. Patricia Bonaldo."				

  

BASE COUNT	147 a	129 c	144 g	134 t	1 others
ORIGIN					

  

Query Match	99.8%; Score 554; DB 9; Length 555;
Best Local Similarity	100.0%; Pred. No. 4.1e-108;
Matches 555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	

  

QY	1	ttttttttcttaataacaatttgacaaaaggtgaagaatcciaaacaaagtattg	60
Db	1	TTTTTTTTCTTTAAATAACAATTGGAAAAAGGCTGAAGAAATCCTTAAGAAGGTATGG	60
QY	61	aggcagatgtccacagcgctcaattcaggtcacagaactgtcccacagacgttcgatcgaact	120
Db	61	AGGCCAGTGCTCCAGGCTCATTCAGTTCCACAGAACTGTCCCTCAGAGACTTCATGGAACCT	120

[illegible]

QY	241	tatatattatataattttacaccaggtacctccagtcactcgttaccattcccaggaga	300
Db	241	TATATATTATATATTATTACATCCAGGTATCCCATCTGTAACATTTCCHAGGGAGA	300
QY	301	catisgctgcctccaaggagacaggaagggttagcgaggaaaggcgacgaagtgc	360
Db	301	CATGGTGTCTTCCAAGGGAGACAGGAAGGGTTAGCGAGGAAAGGGCAGCAGACTGC	360
QY	361	aggctggggttggtgtcaacagaagctgtaggaagcttcagcagctgtagaaggccccgg	420
Db	361	AGGCTGGGGCTTGCTGCTCACAAAGCTGAGAGCTTCAAGCACTGTAAAGAGGCCCCCCGGG	420
QY	421	ctccgcagagacgcaggtlactgaagcaaacgcaatcctccagctccacagccccgccgtsga	480
Db	421	CTCCGCAAGACGCAGAGTACTGAGNMAAGCCAGTCTCTCAAGCTCCACGCCCCCCGCCCTTGCA	480
QY	481	tcacacgcctctctccogaacctcatcatcatcagggcccgcttcatgtcgatccagttg	540
Db	481	TCCACC GCCCTTCTCCGCAAACTCATCATCAAGGCCGCCGTTTATTCGATCCAGTTG	540
QY	541	tgcagcgctgcgcaca	555
Db	541	TGCAGCGCTGCACA	555
RESULT	2		
LOCUS	AM975483	563 bp	mRNA linear EST 02-JUN-2000
DEFINITION	EST387592 MAGE resequences, MAGN Homo sapiens cDNA, mRNA sequence.		

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1.pdb="NCI CGAP GC91"
/tissue_type="germinal center B cell"
/lab_host="DH108"
/note="Vector: PT73D-Pac (Pharmacia) with a modified

```

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 9, 2002, 10:53:56 ; Search time 6963.4 Seconds  
(without alignments)  
1075.740 Million cell updates/sec

Title: US-09-853-544-1

Perfect score: 555  
Sequence: 1 ttttttttctaataac.....agttgtcagcgttcgcaca 555

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: em\_estba:\*  
2: em\_estlum:\*  
3: em\_estlin:\*  
4: em\_estlmu:\*  
5: em\_estlov:\*  
6: em\_estlpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gp\_estl1:\*  
10: gp\_estl2:\*  
11: gp\_hlc:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_inv:\*  
15: em\_gss\_pln:\*  
16: em\_gss\_vrt:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	554	99.8	553	9	AA747315
2	470.6	84.8	565	9	AA747315
3	446.4	80.4	782	10	BF525647
4	444.6	80.1	515	10	BF739767
5	415.6	74.9	422	10	BF436003
6	412.2	74.3	454	9	AI392801
7	402.6	72.5	425	9	AA148983
8	394.4	71.1	455	9	AI697980
9	369.8	66.6	437	9	AA295170
10	364.4	65.7	409	9	AI283647
11	317.8	57.3	343	9	AA687243
12	310.6	56.0	332	9	AI393927
13	265.2	47.8	415	10	HA4044
14	250.2	45.1	782	10	BF754185
15	249	44.9	285	9	AI783873
16	240.2	43.3	400	10	HA4117

18	240	43.2	245	10	R70141	R70141 y148g11.sl
19	213	38.4	248	9	AI918219	AI918219 tn08f10.x
20	178	32.1	580	9	BE110202	BE110202 UI-R-CAO-
21	173.6	31.3	414	10	BF401780	BF401780 UI-R-CAO-
22	168.6	30.4	397	10	BI403166	BI403166 MI-P-CP1-
23	168.6	30.4	504	9	BE114223	BE114223 UI-R-CAO-
24	168.4	30.3	585	10	BF416269	BF416269 UI-R-CAO-
25	167.2	30.1	399	10	BI119093	BI119093 ARO23A09P
26	165.4	29.8	635	10	BF404955	BF404955 UI-R-CA1-
27	160.4	28.9	578	12	AZ488750	AZ488750 IM0319M09
28	156.8	28.3	708	9	BI138537	BI138537 BI138537
29	151	27.2	513	10	BM195829	BM195829 C0317E11-
30	133	24.0	452	9	AI741034	AI741034 w925f08.x
31	90.6	16.3	238	10	BF402467	BF402467 UI-R-CAO-
32	78	14.1	239	10	BE241758	BE241758 TCAPAP2E02
33	78	14.1	282	10	BM193296	BM193296 TCAPAP1E89
34	78	14.1	554	10	BM146800	BM146800 TCAPAP1E73
35	72.2	13.0	753	9	BO098875	BO098875 BO098875
36	69.6	12.5	617	10	BE243742	BE243742 602357051
37	65.8	11.9	622	12	BH305262	BH305262 CH230-4E1
38	65.8	11.9	748	12	BH339028	BH339028 CH230-89F
39	64.4	11.6	740	12	CNS02XF8	AL218285 Tetradon
40	64.2	11.6	692	10	BM388754	BM388754 UI-R-D20-
41	59.2	10.7	298	10	BF151027	BF151027 uy91d04.x
42	58.2	10.5	254	9	AA544572	AA544572 vj67b12.f
43	57	10.3	640	10	BU064111	BU064111 BU064111
44	55.8	10.1	179	9	BA481139	BA481139 BA481139
45	54.2	9.8	243	9	AI597000	AI597000 vj67b12.x

## ALIGNMENTS

RESULT 1	AA747315	555 bp	mRNA	linear	EST 17-MAR-1999
LOCUS	AA747315				
DEFINITION	nx886d02.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1269315 3'				
ACCESSION	AA747315				
VERSION	AA747315.1	GI:2787273			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 555)				
AUTHORS	NCI-CGAP	http://www.ncbi.nlm.nih.gov/ncicgap.			
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index				
JOURNAL	Unpublished (1997)				
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgaabs-r@mail.nih.gov Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman, Ph.D., Gerald Marti, M.D. CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D. CDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.lnl.gov/bhrp/image/image.html Insert Length: 888 Std Error: 0.00 Seq primer: -40m13 fwd. ET from Amerisham High quality sequence stop: 483. Location/Qualifiers				

## FEATURES

source

1..555  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1269315"  
/clone\_lib="NCI\_CGAP\_GCB1"  
/tissue\_type="germinal center B cell"  
/lab\_host="DH10B"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker; Site-1: Not I; Site-2: Eco RI; 1st strand cDNA was prepared from human tonsillar cells enriched for germinal center B cells by flow sorting (CD20<sup>+</sup>, IgD<sup>+</sup>), provided by Dr. Louis M. Staudt (NCI), Dr. David Allman (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was primed with a Not I - oligo(dT) primer  
 5'-TGTACCAATCTGAGTGGAGGCGCCCTCATCTTTT-3'  
 1. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pMT3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonalido."

BASE COUNT 147 a 129 c 144 g 134 t 1 others  
 ORIGIN  
 Query Match 99.8%; Score 554; DB 9; Length 555;  
 Best Local Similarity 100.0%; Pred. No. 4.1e-108;  
 Matches 555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 ttttttttcttaataacatttgcaaaagggtgaagaatccctaacaagatg 60  
 |||||||  
 Db 1 TTTTITTTTCTTAATAAATTTGCAAAAGGTTGAATAATCTTAACAAGTATTG 60  
 Oy 61 aggcacagtgccagagctgcattcagacagaactgtccctcagaagctgtcag 120  
 |||||||  
 Db 61 AGGCCAGTGTCCAGCGTGCATCTCAGTTCACAGAACTGTCTCTCAGGAGTGTGAC 120  
 Oy 121 ggaagaatggtatattacagaagaagaaggagagctgtgcagagaaggagag 180  
 |||||||  
 Db 121 GGAATAATGTATATTATACAGAAAGAAACAGGAGGAGCTTATGTCAGAGAGAGAG 180  
 Oy 181 tgaagcggcacagacatcctctagctcttcattatattatagttatattctata 240  
 |||||||  
 Db 181 TGAAGCGGCAACAGACATCTTACTCTTCAATTATATATGATATATTTCTATA 240  
 Oy 241 tataattatataattatataccaggtatccagtcattcgtacattccaggagaga 300  
 |||||||  
 Db 241 TATATATTTTATATATTTTACATCAGGTATCCAGTATCTGTATCCATTTCCAGAGAGA 300  
 Oy 301 catgggtgctctcagaagcagagaagaagggttagcagaggaaggagcagcagtg 360  
 |||||||  
 Db 301 CATGGGTGCTTCCAGAGCGAGACAGAAAGGTTAGCGAGGAGGAGGCGACGAGTGC 360  
 Oy 361 aggcctggagcttgctcacaagaagctgcagaagcttcagcagctgtgaagagggcccg 420  
 |||||||  
 Db 361 AGGCTGGGGCTTGGCTCACAGAAAGTGCAGAGAGCTTCAGAGCTGTAAAGAGGCGCCCGG 420  
 Oy 421 ctccgcagacagccaggtactgagncaaagccagctcctccagctccagcccgctgcga 480  
 |||||||  
 Db 421 CTCGCGAGAGAGCGCAGTACTGAGNCAAGCCACTCTCCAGCTCCAGCCCGCCCTGCGA 480  
 Oy 481 tgcacggcctctccgaactcattcatatagaagggccggctctatgtgcagcagttg 540  
 |||||||  
 Db 481 TGCACGCGCTTCTCCGAAACTTCATCATCATGAGGCGCGCTTCATGTGATCCAGTTG 540  
 Oy 541 tgcacgctgcgcaca 555  
 |||||||  
 Db 541 TGCACGCTGCGCACA 555  
 RESULT 2  
 AM975483 563 bp mRNA linear EST 02-JUN-2000  
 LOCUS EST387592 MAGC resequences, MAGN Homo sapiens cDNA, mRNA sequence.  
 DEFINITION AM975483  
 ACCESSION AM975483  
 VERSION AM975483.1 GI:8166697  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE 1 (bases 1 to 563)

AUTHORS Hegde, P., Qi, R., Abernathy, K., Dharap, S., Gaspard, R., Gay, C., Holt, I. E., Saeed, A. I., Sharov, V., Lee, N. H., Yeatman, T. J. and Quackenbush, J.  
 TITLE Assessment of gene expression patterns in a model of colon tumor metastasis using a 19,200 element cDNA microarray  
 JOURNAL Unpublished (2000)  
 COMMENT The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 3528  
 Fax: 301 838 0208  
 Email: johnh@tigr.org  
 Plate: 349  
 Seq primer: Forward.  
 FEATURES  
 source Location/Qualifiers  
 1..563  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="MAGE resequences, MAGN"  
 /note="Vector: pBluescriptSkm"  
 BASE COUNT 152 a 131 c 144 g 136 t  
 ORIGIN  
 Query Match 84.8%; Score 470.6; DB 9; Length 563;  
 Best Local Similarity 94.9%; Pred. No. 2.5e-90;  
 Matches 507; Conservative 0; Mismatches 25; Indels 2; Gaps 2;  
 Oy 11 ttttaatacaatttgacaaaagggtgaagaatccctaacaaggtatggaagcagtg 70  
 |||||||  
 Db 1 TTTAAATACATTGACAAAAGGTTGAAGAAATCTTAACAAGATTTGAGGCGCAGTGT 60  
 Oy 71 ccaggtcgtcattcagttccacagaactgtcctcagagcgttcgcatggaactggaaatgt 130  
 |||||||  
 Db 61 CCAGGCTGATTCAGTTCACAGACAGTCTCTCAGACAGTGTGCATGGAATGTGT 120  
 Oy 131 ataattacagaagaacacaggaaggacttagtgcagagaaggagaagcagtggtgaaggagca 190  
 |||||||  
 Db 121 ATATATTCAGAAAGAAACAGGAGGAGCTTATGTCACAGAGAGACTAGTGTGACGCGCA 180  
 Oy 191 aagagctccttgctcttataattatataatgatatgatatcttatattctatatataatta 250  
 |||||||  
 Db 181 ACAGCATCTTGTGCTTCTTATTTATATATGATATATTTCTTATATATATTTTA 240  
 Oy 251 tataattatatacaattatccagtcattcgttacatttccacagggagagacatgggtgct 310  
 |||||||  
 Db 241 TATATTTGAACATCCAGATATCCAGATCATCTGTACATTTCACAGGAGACATATGGGTCT 300  
 Oy 311 tccaaggcagacaggaagaagggttagcaggaaggagcagcagcgtgcagcgtgcg 370  
 |||||||  
 Db 301 TCCAAAGCGAGACAGAGAAAGGTTAGCGAGGAGGCGACGCGTCCAGGCTTCGCGC 360  
 Oy 371 ttggctcacaagaagctgcaggaagcttcacagcgtgtaagaaggcccgagctccgcagac 430  
 |||||||  
 Db 361 TTGGCTCACAGAGAGCTGCAGAGCTTCACAGATGTAAAGGGCCCCGGGCTTCCGACAGC 420  
 Oy 431 gccaggtact-gagncaaaagccagttcctcagcctccacagcccgctcgtatccacgcgc 489  
 |||||||  
 Db 421 GCCAGTACTGGCAGCAAAAGCCAGTCTCCAGCTCCAGCCCGCCGCTGGAAACCAAGCGCC 480  
 Oy 490 ttctccgcaaaacttcataca-tcagggcccgctcattcgtcagatccagttgtg 542  
 |||||||  
 Db 481 TTGTTTCGCAAACTTATTTATTTTCAGGGCCGCTTATGTTCAATCCATTTTTG 534  
 RESULT 3  
 BF525647 782 bp mRNA linear EST 11-DEC-2000  
 LOCUS BF525647  
 DEFINITION 602069764P1 NCI\_CGAP\_Brn64 Homo sapiens cDNA clone IMAGE:4212734  
 5', RNA sequence.  
 ACCESSION BF525647  
 VERSION BF525647.1 GI:11613008  
 KEYWORDS EST.

SOURCE	ORGANISM	human.
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	
AUTHORS	1 (bases 1 to 782)	
TITLE	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .	
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC) unpublished (1999)	
COMMENT	Contact: Robert Strausberg, Ph.D. Email: <a href="mailto:cgabbs-remail.nih.gov">cgabbs-remail.nih.gov</a> Tissue Procurement: David N. Louis, M.D. cDNA Library Preparation: Life Technologies, Inc. DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <a href="http://image.lnl.gov">http://image.lnl.gov</a> Plate: L1AM9783 row: P column: 15 High quality sequence size: 599.	
FEATURES	location/Qualifiers	
SOURCE	1..782	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/clone="IMAGE:4212734"	
	/clone_lib="NCI CGAP Brn64"	
	/tissue_type="gliblastoma with EGFR amplification"	
	/lab_host="DH10B (T1 phage-resistant)"	
	/note="Organ: brain; Vector: pCMV-SPORT6; Site: 1: NotI; Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.57 kb. Constructed by Life Technologies. Note: this is a NCI CGAP Library."	
BASE COUNT	186 a 216 c 204 g 174 t 2 others	
ORIGIN		
Query Match	80.4%; Score 446.4; DB 10; Length 782;	
Best Local Similarity	94.9%; Pred. No. 3.6e-85;	
Matches	525; Conservative 0; Mismatches 22; Indels 6; Gaps	6
OY	1 ttttttctttaataacaatttgacaaaagggtgaagaatccctaacaaggtatg 60	
Db	713 TTTTTCGTCGTATTCACAAATGACAAAGGAGAAATTCCT-AACAAGTATG 655	
OY	61 aggcagagtcacagagctgcattcattacacagaactgtctccaagagcttga-tgaaac 119	
Db	654 AGCCAGAGTGCCA-CTGTGATTCAGTTTCAACAACGTGCTCAGGACGTTGATGGAAc 596	
OY	120 tggaaatgtgataaatacagaagaagaacagggagacttattgcagagagagagact 179	
Db	595 TGGAAATCGGTATATTATTCAGAAAG-TACCAGGAGGAGACTTATGTCAGAGAGAGAGACT 537	
OY	180 gtgagcgggcaacagcatccttaagct-ttcaattattatatagtgtatatttcta 238	
Db	536 GTGGACGGGCAACACATCCCTTAAGTCTGTTCAATTTATATATGATATGATTTTCTA 477	
OY	239 tataataattataataattttacatccaggtatcccaagtcattgtacacattcccaaggga 298	
Db	476 TATATATATTTATATATTTTACATCCAGGTATCCAGTCAATCTGAAACCATTTTCCAGGGA 417	
OY	299 gacatgggtgtctcccaaggcagagcagaagaaggttaaggcagaggaaggcagcagcgt 358	
Db	416 GACATGGGTGCTCCCAAGGCGAGACAGAAAGGTTAAGCAGAGGAAAGGGGCAAGCGAGGT 357	
OY	359 gcaagcttgaggcttgcctacagagaagcttgcagagacttcaagcagctgtlaaaggggcccg 418	
Db	356 GCAGCGCTGGGGCTTGCGTCACAGAAAGCTGCAGAGACTTACGAGCACTGTAAGAGGGCCCG 297	
OY	419 ggcctcgcaagcagcaggtact-gagncaaagccagtcctccagctccagcccgctg 477	
Db	296 GGCTTCGAGAGAGCGAGGACTGCGAGCAAAACCAAGTCTCCAGGCTCCACGCCGCCGCTG 237	
OY	478 cgatcacacgctcttcgcaaatcatcatcatcagagggcccgcttcatgtcagaccg 537	

Db	236	CGATCCACCGGCTCTCCGCCAACAATCTACTATCATTACAGAGGCGCCGCTTATGTCATCCAG	177
QY	538	ttgtgcagcgtcgc 550	
Db	176	TTGTGCAGCGTGC 164	
RESULT	4		
LOCUS	BF739767		
DEFINITION	h559b11.x1 NCI_CGAP_Brn41 Homo sapiens cDNA clone IMAGE:3174117 3',		
ACCESSION	BF739767		
VERSION	BF739767.1		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 515)		
JOURNAL	NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.		
COMMENT	National Cancer Institute / National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project (CGAP/BTCAP), Tumor Gene Index		
UNPUBLISHED	(1998)		
CONTACT	Robert Strausberg, Ph.D.		
EMAIL	cgapbs-remail.nih.gov		
TISSUE	Procurement: Robert Jenkins, M.D.		
CDNA	Library Preparation: M. Bento Soares, Ph.D. and M. Fatima		
CDNA	Library Arrayed by: The I.M.A.G.E. Consortium/LLNL		
CDNA	Sequencing by: Washington-University Genome Sequencing Center		
CDNA	Distribution: NCI-CGAP clone distribution information can be		
CDNA	found through the I.M.A.G.E. Consortium/LLNL, send email to:		
CDNA	info@image.llnl.gov		
CDNA	Seq primer: -400p from Gidco		
CDNA	High quality sequence stop: 462.		
FEATURES			
SOURCE	Location/Qualifiers		
	1..515		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:3174117"		
	/clone_lib="NCI CGAP_Brn41"		
	/tissue_type="Oligodendrogloma"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: brain; Vector: p773D-Pac (Pharmacia) with a		
	modified polylinker; Site.1: NotI; Site.2: EcoRI; 1st		
	strand cDNA was primed with a Not I - oligo(dT) primer [5'		
	TGTACCAATCTGAAATGGAGGAGCGCGCATCTTTTTTTTTTTTTTT 3'];		
	double-stranded cDNA was ligated to Eco RI adaptors		
	(Pharmacia), digested with Not I and cloned into the Not		
	I and Eco RI sites of the modified p773 vector. Library		
	constructed and normalized by Bento Soares and M. Fatima		
	Bonaldo."		
BASE COUNT	153 a 105 c 125 g 132 t		
ORIGIN			
Query Match	80.1%; Score 444.6; DB 10; Length 515;		
Best Local Similarity	97.7%; Pred. No. 8.6e-85;		
Matches	461; Conservative 0; Mismatches 10; Indels 1; Gaps 1;		
QY	1	tttttttcttaataacaatttgacaaaaaggtgaagaatcctaacaaggtatg 60	
Db	44	TTTTTTTCTTTTAATAAATTTGACAAAAGGCTGAAAATTCCTAACAAGTATG 103	
QY	61	agggcagtgctcagcgtcgtcattcagttcacagaactgtccctcaggagcgttgatggaact 120	
Db	104	AGGCCAGTGTCCAGGCTCATTCATTCAACAAAAGTCTCTCAGACGCTTGATGAGACT 163	
QY	121	ggaatgtgtataattacaagaagaaagagagagacttagtgcagagagagagagagactg 180	
Db	164	GGAATGTGTATAATTACAGAAAAACAGGAGGAGACTTATGTCAAAAGAGAAAGAGAGTG 223	

QY 181 tggacgggcacacagcatccttagcttccatattatataatggtatatttctata 240  
|||||  
Db 224 TGACAGGGGACAGACATTCCTTACTCTTTCATATTATATATGATATATTTCTATA 263  
QY 241 tatatttatattatattacatccaggtatccagtcacatctgtacattccaggaga 300  
|||||  
Db 284 TATATATTTATATATTTTATCATCCAGGTATCCAGTATCTGTACCATTTCCAGAGAGA 343  
QY 301 catgagtgcttccaaggacagagaagggttagcagggagggagggagcagcgttc 360  
|||||  
Db 344 CATGGGTGCTCCAGAGCGGACAGAGAAAGGTTAGCGAGGAGGCGACGACGTCG 403  
QY 361 aggcctgggcttgcctcacagaagctgcagagcttcagcagctgtgaagagggcccg 420  
404 AGCGTGGGGGCTTGCTCACAGAGCTGCAGAGACTTCAAGGACTGTAAAGGGCCCGG 463  
QY 421 ctccgacagagcaggtact-gagncagaagccagctcctccagctccagccc 471  
|||||  
Db 464 CTCGCGAGAGCGCAGGTACTGCGACGAAAGCCAGTCTCTCAGCTCCAGCCGCC 515

RESULT 5  
BF436003 422 bp mRNA linear EST 19-JAN-2001  
LOCUS nab32c01.x1 Soares\_NSF\_F8\_9W\_OT\_PA\_P\_S1 Homo sapiens cDNA clone  
DEFINITION IMAGE:3267456 3', mRNA sequence.  
ACCESSION BF436003  
VERSION BF436003.1 GI:11448318  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 422)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-rt@mail.nih.gov  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: -40UP from Gibco  
High quality sequence stop: 418.  
Location/Qualifiers

## FEATURES

source

1. 422  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3267456"  
/clone\_lib="Soares\_NSF\_F8\_9W\_OT\_PA\_P\_S1"  
/lab\_host="DH10B"  
/note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with  
a modified polylinker; Site\_1: Not I; Site\_2: Eco RI;  
Equal amounts of plasmid DNA from five normalized  
libraries were mixed, and ss circles were made in vitro.  
Following HAP purification, this DNA was used as tracer in  
a subtractive hybridization reaction. The driver was  
PCR-amplified cDNAs from pools of 5,000 clones made from  
the same 5 libraries. The pools consisted of the following  
libraries and clones: Soares NBHSF pool 1:  
309384-310919, 323208-325895 Soares NB2HP pool 1:  
145032-147335, 147720-148103, 148872-149255, 15002 -  
150407, 151176-152327 Soares NB2HF8-9W pool 1:  
758280-760583, 772104-774407 Soares NBHFA pool 1:  
304776-306311, 320136-322823, 326280-326653 Soares NBHOT  
pool 1: 723720-726407, 739080-740959 Subtraction by Bento  
Soares and M. Fatima Bonaldi."

BASE COUNT  
ORIGIN

121 a 80 c 115 g 106 t

Query Match

74.9%; Score 415.6; DB 10; Length 422;

Best Local Similarity 99.1%; Pred. No. 1,3e-78;  
Matches 418; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 tttttttaaataaacaatttgacaaaagggggaagaaatccaaacaggtattgagc 64  
|||||  
Db 1 TTTTCTTTTAAATTAACAAATTTGACAAAGGGTGAAGAAATCCATAACAGGTATTTAGGC 60  
QY 65 cagtgccagggcgcgtcattcagttcacagacgtgctccaggagcgttcgtgaactgga 124  
61 CATGTGCCAGGCTGCAGTTTCAGTTCACAGAACTGCTCCAGAGCTTGCAATGGAACTGGAA 120  
QY 125 atgtgtataattacagaagaataacaggagagcttagtcagagaggagacaggtgtga 184  
121 ATGTGTATTAATTAACAGAAATAAACAGGAGGAGCTTGTGTGTCAGAGAGACTATGTGTGA 180  
QY 185 cgggacaagcagcctccttagcttcttatattatattatgtatatttccatataata 244  
|||||  
Db 181 CGGCAACAGCATCCCTTAATCTTTCAATTTATATATGATATATTTCTATATATA 240  
QY 245 tatatatattatattacatccaggtatccagtcacatctgtacattccaggagaatg 304  
|||  
Db 241 TATCTATATATTTTACATCCAGGTATCCAGTATCTGTACATTTCCAGGAGAGCATG 300  
QY 305 ggtgtctccaaggcagagacagaagaagggttaaggcagggagggcgacgagctcagc 364  
|||||  
Db 301 GGTGCTTCCAGGCGAGACAGGAAAGGTTAGGCAAGGAGGCGACGATGTCAGGC 360  
QY 365 tggggcttgctccacagaagctgcagagacttcagagctgtgaagaaggcccgagctcc 424  
|||||  
Db 361 TGGGCTTGCTCACAGAAAGCTGCAGAGCTTCAGCACTGTAAAGGGCCCGGCTCC 420  
QY 425 gc 426  
||  
Db 421 GC 422

RESULT 6  
A1392901 454 bp mRNA linear EST 30-MAR-1999  
LOCUS tg10e07.x1 NCI\_CGAP\_CLL1 Homo sapiens cDNA clone IMAGE:2108388 3',  
DEFINITION mRNA sequence.  
ACCESSION A1392901  
VERSION A1392901.1 GI:4222448  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 454)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-rt@mail.nih.gov  
Tissue Procurement: Ash Alizadeh, John Byrd, M.D., Mike Grever,  
M.D., Louis M. Staudt, M.D., Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D.  
DNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution Information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
www.bio.llnl.gov/dbfp/image/image.html  
Insert length: 1175 Std Error: 0.00  
Seq primer: -40UP from Gibco  
High quality sequence stop: 444.  
Location/Qualifiers

## FEATURES

source

1. 454  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2108388"  
/clone\_lib="NCI\_CGAP\_CLL1"  
/tissue\_type="B-cell, chronic lymphocytic leukemia"





TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
CONTACT: Robert Strausberg, Ph.D.  
COMMENT Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
Emmett-Buck, M.D., Ph.D.  
CDNA Library Preparation: M. Bento Soares, Ph.D.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www.bio.llnl.gov/dbp/image/image.html  
Insert Length: 615 Std Error: 0.00  
Seq primer: -40UP from Gibco.  
Location/Qualifiers  
1. 455  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2341562"  
/clone\_lib="NCI-CGAP\_Lu24"  
/tissue\_type="carcinoid"  
/lab\_host="DH10B"  
/note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a  
modified polylinker; Plasmid DNA from the normalized  
library NCI-CGAP\_Lu5 was prepared, and ss circles were  
made in vitro. Following HAP purification, this DNA was  
used as tracer in a subtractive hybridization reaction.  
The driver was PCR-amplified cDNAs from a pool of 5,000  
clones made from the same library (clonoids  
1414920-1417991 and 1520904-1522439). Subtraction by Bento  
Soares and M. Fatima Bonaldo."

BASE COUNT 136 a 80 c 108 g 130 t 1 others  
ORIGIN

Query Match 71.1%; Score 394.4; DB 9; Length 455;  
Best Local Similarity 98.1%; Pred. No. 4.3e-74;  
Matches 409; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 1 ttttttttcttaataacatttgacaaagggtgaagaatccttaacaagttatg 60  
|||||  
Db 40 ttttttttcttaataacatttgacaaagggtgaagaatccttaacaagttatg 99  
|||||

QY 61 aggcagtgctccaggtcgaatcagtcacagaactgctccagagcgttcaggaact 120  
|||||  
Db 100 AGGCAGTGTCCAGGCTGCACTTCAAGTTCACAAAGCTGCTCCTCGAGCGTTGCGACT 159  
|||||

QY 121 ggaatggttatataacagaagaaggagagagctagtcagagaagagagagtg 180  
|||||  
Db 160 GGAATGTGTATATATACAGAAAAACAGGAGGAGCTTAGTGCANAGAGAGACAGTG 219  
|||||

QY 181 tgaaggggaacagcacccttagcttctcatattatataatgatatatttctata 240  
|||||  
Db 220 TGAGCGGCGACAGCATCTTACTCTTATATATATATATATATATATATATATAT 279  
|||||

QY 241 tatatatatatattatatacaccaggtatccagtcacatcgtaccattccaggagaga 300  
|||||  
Db 280 TAT 338  
|||||

QY 301 catggtgcttccaagggcagaaggaagggttaaggcaggaaggagcagcagtg 360  
|||||  
Db 339 CATGGGTGCTTCCAAAGGCGAGAGAAAGGTTAAGGAGGAGGAGGAGCGACGCTGC 398  
|||||

QY 361 aggcctggagcttgctcacagaagctgcaggaagcttcacagcgttcaagagggccc 417  
|||||  
Db 399 AGGCTGGGCTTGGCTCAGAGAACTGCGAGAGCTTACAGCACTGTAAAGAGGGCCCC 455  
|||||

RESULT 9  
AM295170 437 bp mRNA linear EST 16-JAN-2000  
LOCUS  
DEFINITION UI-H-B12-1b-d-03-0-UI.s1 NCI-CGAP\_Sub4 Homo sapiens cDNA clone

IMAGE:2728733 3', mRNA sequence.  
AM295170  
AM295170.1 GI:6701806  
EST.  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
1 (bases 1 to 437)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
CONTACT: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
The sequence contained an oligo-dT track that was present in the  
oligonucleotide that was used to prime the synthesis of first  
strand cDNA and therefore this may represent a bonafide poly A  
tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution:  
NCI-CGAP clone distribution information can be found through the  
I.M.A.G.E. Consortium/LLNL at:  
www.bio.llnl.gov/dbp/image/image.html The following repetitive  
elements were found in this cDNA sequence: 269-314, >(TA  
)n\$SimpleRepeat  
Seq primer: M13 Forward  
POLYA=yes.  
Location/Qualifiers  
1. 437  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2728733"  
/clone\_lib="NCI-CGAP\_Sub4"  
/lab\_host="DH10B (Life Technologies)"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker; Site 1: Not I; Site 2: Eco RI; The  
NCI-CGAP\_Sub4 library is a subtracted library derived from  
the NCI-CGAP\_Sub2 library which is a subtracted library  
derived from the NCI-CGAP\_Sub1 library, which is a  
mixture of 21 normalized or subtracted NCI-CGAP  
libraries: NCI-CGAP\_C04, NCI-CGAP\_Pr22, NCI-CGAP\_Pr28,  
NCI-CGAP\_C010, NCI-CGAP\_C016, NCI-CGAP\_Kid5,  
NCI-CGAP\_Kid12, NCI-CGAP\_Kid3, NCI-CGAP\_Kid11,  
NCI-CGAP\_Lym2, NCI-CGAP\_Br2, NCI-CGAP\_C08, NCI-CGAP\_C11L1,  
NCI-CGAP\_Lu24, NCI-CGAP\_Br23, NCI-CGAP\_Lu5,  
NCI-CGAP\_Lu24, NCI-CGAP\_Lu19, NCI-CGAP\_GC4, NCI-CGAP\_GC6,  
NCI-CGAP\_Br25. These 21 libraries were pooled and a  
single-stranded DNA preparation of the resulting mixture  
was used as a tracer in a subtractive hybridization with  
a driver whose composition is detailed below:  
NCI-CGAP\_Kid3 pool 1 : LHAM 3334-3337, 3682-3683,  
3798-3803 (IMAGE Clonoids 1323376-1323911,  
1456008-1456775, 1500552-1502855) NCI-CGAP\_Kid5 pool 1 :  
LHAM 3338-3342, 3722-3725, 3776-3778 (IMAGE Clonoids  
1323912-1325831, 1471368-1472903, 1492104-1493255)  
NCI-CGAP\_Lu5 pool 1 : LHAM 3575-3582, 3851-3854 (IMAGE  
Clonoids 1414920-1417991, 1520904-1522439) NCI-CGAP\_GC4  
pool 1 : LHAM 3164-3167, 3716-3720, 3733-3735 (IMAGE  
Clonoids 1257096-1258631, 1469064-1470983, 1475592-1476743  
) NCI-CGAP\_Pr22 pool 1 : LHAM 2457-2459, 2758-2759,  
3062-3068 (IMAGE Clonoids 985608-986759, 1101199-1101959,  
1217928-1220615) NCI-CGAP\_GC4 pool 1 : LHAM 2644-2653,  
2871-2872 (IMAGE Clonoids 105716-1061255, 1144584-1145351)  
) Subtraction was performed as previously described  
(Bonaldo, Lennon & Soares (1996): Normalization and  
Subtraction: Two Approaches To Facilitate Gene Discovery.  
Genome Research 6, 791-806.)  
TAG\_Lib=NCI-CGAP\_GC4  
TAG\_Tissue=germ cell  
TAG\_SEQ=AAATC"

BASE COUNT 126 a 73 c 102 g 136 t  
ORIGIN

Query Match 66.6%; Score 369.8; DB 9; Length 437;  
Best Local Similarity 99.5%; Pred. No. 7.6e-69;  
Matches 371; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ttttttttcttaataaattgacaaaaggtgagaatactctaacaagatattg 60  
|||||  
DB 60 ttttttttcttaataaattgacaaaaggtgagaatactctaacaagatattg 119  
|||||

QY 61 aggcagtgccagcagtcagttcagttcaagaactgtccctcagagcgttcagtaact 120  
|||||  
DB 120 AGGCCAGTGTCCAGCGCTGCATTTCAGTTCACAGACATGTCTCAGGACGTTCATGCACT 179  
|||||

QY 121 ggaatgtctataataacagaagaacacagggagagacttagtgcagagagagacagtg 180  
|||||  
DB 180 GGAATGTGTATTAATTAACAAGAAAAACAGGAGGAGACTTAGTGCAGAGACAGCAGTG 239  
|||||

QY 181 tggacgggcaacagcagtcctccttagtcttcattatattatattatattatattatattat 240  
|||||  
DB 240 TGGACGGGCAACAGCATCCTTACGTCTTCAATATTATTAATTAATTAATTAATTAATTAATTA 299  
|||||

QY 241 tatattatattatattatattatattatattatattatattatattatattatattatattat 300  
|||||  
DB 300 TATATATTATATTATTTTACATCCAGGTATCCAGTCATCTGTACCATTTTCCACAGGAGACA 359  
|||||

QY 301 catgggtgcttccaagcgagacagaaaggttagcagagagagggcgagcgagtg 360  
|||||  
DB 360 CATGGGTGCTTCCAGGCGAGACAGAAAGGTTAGCAGGAGAGGCGACGCGTG 419  
|||||

QY 361 aggcgtgggcttg 373  
|||||

DB 420 AGGCTGGGCTCG 432  
|||||

RESULT 10  
AI283647 409 bp mRNA linear EST 28-JUN-1999  
LOCUS qj67d05.x1 NCL\_CGAP\_Kid3 Homo sapiens cDNA clone IMAGE:1864521 3',  
DEFINITION mRNA sequence.  
ACCESSION AI283647  
VERSION AI283647.1 GI:3921880  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 409)  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-r@mail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: M. Bento Soares, Ph.D.  
DNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www.bio.lnl.gov/bbrp/image/image.html](http://www.bio.lnl.gov/bbrp/image/image.html)  
Insert Length: 725 Std Error: 0.00  
Seq primer: -40UP from Glibco  
High quality sequence stop: 408.  
Location/Qualifiers  
1. 409  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_1ib="IMAGE:1864521"  
/clone\_1ib="NCL\_CGAP\_Kid3"  
/lab\_host="PH10B"  
/note="Organ: kidney; Vector: pRT3D-Pac (Pharmacia) with  
a modified polylinker. Site\_1: Not I; Site\_2: Eco RI; 1st

strand cDNA was primed with a Not I - Oligo(dT) primer,  
double-stranded cDNA was ligated to Eco RI adaptors  
(pharmacia), digested with Not I and cloned into the Not  
I and Eco RI sites of the modified pRT3 vector. mRNA  
source: 2 pooled kidneys. Library went through one round  
of normalization. Library constructed by Bento Soares and  
M. Fatima Bonaldo.

BASE COUNT 125 a 68 c 97 g 118 t 1 others  
ORIGIN

Query Match 65.7%; Score 364.4; DB 9; Length 409;  
Best Local Similarity 99.5%; Pred. No. 1.1e-67;  
Matches 365; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ttttttttcttaataaattgacaaaaggtgagaatactctaacaagatattg 60  
|||||  
DB 43 ttttttttcttaataaattgacaaaaggtgagaatactctaacaagatattg 102  
|||||

QY 61 aggcagtgccagcagtcagttcagttcagagactgtccctcagagcgttcagtaact 120  
|||||  
DB 103 AGGCCAGTGTCCAGCGCTGCATTTCAGTTCACAGACATGTCTCAGGACGTTCATGCACT 162  
|||||

QY 121 ggaatgtctataataacagaagaacacagggagagacttagtgcagagagagacagtg 180  
|||||  
DB 163 GGAATGTGTATTAATTAACAAGAAAAACAGGAGGAGACTTAGTGCAGAGACAGCAGTG 222  
|||||

QY 181 tggacgggcaacagcagtcctccttagtcttcattatattatattatattatattatattatattat 240  
|||||  
DB 223 TGGACGGGCAACAGCATCCTTACGTCTTCAATATTATTAATTAATTAATTAATTAATTAATTA 282  
|||||

QY 241 tatattatattatattatattatattatattatattatattatattatattatattatattat 300  
|||||  
DB 283 TATATATTATATTATTTTACATCCAGGTATCCAGTCATCTGTACCATTTTCCACAGGAGACA 342  
|||||

QY 301 catgggtgcttccaagcgagacagaaaggttagcagagagagggcgagcgagtg 360  
|||||  
DB 343 CATGGGTGCTTCCAGGCGAGACAGAAAGGTTAGCAGGAGAGGCGACGCGTG 402  
|||||

QY 361 aggcgtgggcttg 367  
|||||

DB 403 AGGCTTG 409  
|||||

RESULT 11  
AA687243 343 bp mRNA linear EST 24-DEC-1997  
LOCUS nv60g02.s1 NCL\_CGAP\_GCB1 Homo sapiens cDNA clone IMAGE:1234226 3',  
DEFINITION mRNA sequence.  
ACCESSION AA687243  
VERSION AA687243.1 GI:2675434  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 343)  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-r@mail.nih.gov  
unknown library type  
Insert Length: 921 Std Error: 0.00  
Seq primer: -40m13 fwd. ET from Amersham  
High quality sequence stop: 329.  
Location/Qualifiers  
1. 343  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1234226"  
/clone\_1ib="NCL\_CGAP\_GCB1"





RESULT 15  
R70140/c  
LOCUS  
DEFINITION Y148p11\_r1 Soares placenta Nb2HP Homo sapiens cDNA clone  
ACCESSION R70140  
VERSION R70140.1 GI:843657  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 432)  
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman  
M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,  
Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston  
R., Williamson,A., Woldmann,P. and Wilson,R.  
The WashU-Merck EST Project  
Unpublished (1995)  
CONTACT: Wilson RK  
JOURNAL Washington University School of Medicine  
COMMENT 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: estewatson.wustl.edu  
Insert Size: 650  
High quality sequence stops: 309  
Source: IMAGE Consortium, LNL  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.lnl.gov) for further information.  
Insert Length: 650 Std Error: 0.00  
Seq primer: M13Rpl  
High quality sequence stop: 309.

## FEATURES

source  
1..432  
location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="GDB:515192"  
/db\_xref="taxon:9606"  
/clone="IMAGE:142532"  
/clone.lib="Soares placenta Nb2HP"  
/sex="Female"  
/dev\_stage="placenta obtained at birth (full term)"  
/lab\_host="DH10B (ampicillin resistant)"  
/note="Organ: Placenta; Vector: p77Y3D (Pharmacia) with a  
modified polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st  
strand cDNA was primed with a Not I - oligo(dT) primer [5'  
AACTGGAGAAATTCGCGCCGCGAGAGATTTTCTTTTCTTTT 3']  
double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified p77Y3 vector. Library  
went through one round of normalization. Library  
constructed by Bento Soares and M. Fatima Bonaldo. "

## BASE COUNT

95 a 110 c 125 g 100 t 2 others

## Query Match

Best Local Similarity 45.1%; Score 250.2; DB 10; Length 432;

Matches 309; Conservative 0; Mismatches 46; Indels 3; Gaps 3;

QY 200 ttatgtcttcataattatattatgtatatttctatatatatatatatttata 259  
Db 393 TCATATTTTAAATVATNGTAATVGTAAATTTCTATATNATATATTCATATTTTAAAC 334  
QY 260 catccagatcatccagatcatcgtaccatt-cccgaggagacatggtgtccaaaggc 318  
Db 333 ATCCCGATATCCCATCTCTGTACCATTTTCCCGAGGAGACATGGTGCTTCCAAGGC 274  
QY 319 gagacagaaaggttaggcagggaaagggcagcaggtgcaggtgtggtggtgtgctca 378  
Db 273 GAGACAGAAAGGTTAGCAGGAGGAGGCGACGACGCTGCGGCTTGCGCTCA 214

QY 379 cagaagctgcaggagcttcagcagctgtlaagaggccccggggtccgcagacgccaagta 438  
Db 213 CAGAGCTGCAGAGAGCTTACAGCGACGTAGAGAGGCCCCGGG-TCGCGACAGCGCAGGTA 155  
QY 439 ct-gagncaaagccagtcctccagctccagctccagcccgctgcagtcacagcgtctccgc 497  
Db 154 CTGGCAGCAAAAGCCAGTCTTCAGAGCTCCACGCCCGCTCGATCCACCCCTTCTCCGC 95  
QY 498 aaactcatcatcatcagggcccgcttcagtcagtcagcttgagcaggtgagca 555  
Db 94 AAACCTCATCATCATCAGGCGCCGCTTCATGTGATTCAGTGTGCAAGCTGCGACAA 37

Search completed: June 9, 2002, 17:27:05  
Job time: 23589 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 9, 2002, 15:17:46 ; Search time 190.69 Seconds  
(without alignments)  
714.912 Million cell updates/sec

Title: US-09-853-544-1

Perfect score: 555  
Sequence: 1 ttttttttttctaataac.....agtgtgcagcgtgcgaca 555

Scoring table:  
IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Issued Patents, NA: \*  
1: /cgn2\_6/ptodata/2/1na/5A.COMB.seq: \*  
2: /cgn2\_6/ptodata/2/1na/5B.COMB.seq: \*  
3: /cgn2\_6/ptodata/2/1na/6A.COMB.seq: \*  
4: /cgn2\_6/ptodata/2/1na/6B.COMB.seq: \*  
5: /cgn2\_6/ptodata/2/1na/PCRTUS.COMB.seq: \*  
6: /cgn2\_6/ptodata/2/1na/Backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41.6	7.5	7218	1	Sequence 14, Appl
2	38.8	7.0	2875	5	Sequence 63, Appl
3	36.8	6.6	71989	4	Sequence 2, Appl
4	35.8	6.5	3777	3	Sequence 15, Appl
5	35.8	6.5	3777	4	Sequence 15, Appl
6	35.2	6.3	6769	1	Sequence 20, Appl
7	35.2	6.3	6769	1	Sequence 20, Appl
8	35.2	6.3	6769	1	Sequence 20, Appl
9	35.2	6.3	6769	1	Sequence 20, Appl
10	35.2	6.3	6769	1	Sequence 20, Appl
11	35.2	6.3	6769	4	Sequence 20, Appl
12	35.2	6.3	6769	5	Sequence 20, Appl
13	35.2	6.3	6769	5	Sequence 20, Appl
14	35.2	6.3	6769	5	Sequence 20, Appl
15	35.2	6.3	6769	3	Sequence 1, Appl
16	35.2	6.3	6769	4	Sequence 1, Appl
17	35.2	6.3	6769	4	Sequence 1, Appl
18	35.2	6.3	6769	4	Sequence 1, Appl
19	35.2	6.3	6769	4	Sequence 1, Appl
20	35.2	6.3	6769	4	Sequence 1, Appl
21	34.2	6.2	1420	1	Sequence 1, Appl
22	34.2	6.2	2689	2	Sequence 1, Appl
23	34.2	6.2	2689	3	Sequence 1, Appl
24	34.2	6.2	2689	3	Sequence 1, Appl
25	33.8	6.1	600	4	Sequence 1, Appl
26	33.8	6.1	2028	4	Sequence 1, Appl
27	33.8	6.1	2734	3	Sequence 79, Appl

28	33.8	6.1	2821	4	US-09-135-021A-115	Sequence 115, App
29	33.8	6.1	3181	3	US-09-135-021-1	Sequence 1, Appl
30	33.8	6.1	3181	4	US-09-135-020-1	Sequence 1, Appl
31	33.8	6.1	3181	4	US-09-135-010A-1	Sequence 1, Appl
32	33.8	6.1	3182	3	US-09-135-021-5	Sequence 5, Appl
33	33.6	6.1	5526	4	US-08-751-359-21	Sequence 21, Appl
34	33.6	6.1	5526	4	US-08-907-146-21	Sequence 21, Appl
35	33.2	6.0	1789	1	US-08-455-543A-29	Sequence 29, Appl
36	33.2	6.0	1789	1	US-08-223-305C-29	Sequence 29, Appl
37	33.2	6.0	2338	1	US-08-455-543A-31	Sequence 31, Appl
38	33.2	6.0	2338	2	US-08-223-305C-31	Sequence 31, Appl
39	33.2	6.0	5904	1	US-07-745-206A-6	Sequence 6, Appl
40	33.2	6.0	5904	1	US-08-455-543A-3	Sequence 3, Appl
41	33.2	6.0	5904	2	US-08-193-078B-3	Sequence 3, Appl
42	33.2	6.0	5904	2	US-08-223-305C-3	Sequence 3, Appl
43	33.2	6.0	5904	2	US-08-149-097D-3	Sequence 3, Appl
44	33.2	6.0	5904	2	US-08-311-363-6	Sequence 6, Appl
45	33.2	6.0	6575	3	US-08-949-386-3	Sequence 3, Appl

#### ALIGNMENTS

Result 1  
US-08-232-463-14/C  
Sequence 14, Application US/08232463  
Patent No. 5670367  
GENERAL INFORMATION:  
APPLICANT: DORNER, F.  
APPLICANT: SCHEIFLINGER, F.  
APPLICANT: FALKNER, F. G.  
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 1800 Diagonal Road, Suite 500  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22313-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/232,463  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/935,313  
FILING DATE:  
APPLICATION NUMBER: EP 91 114 300.6  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 304722/114 IMMU  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)836-9300  
TELEFAX: (703)683-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7218 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
CLONE: pTZgpt-Fls  
US-08-232-463-14

	Query Match	7.5%:	Score 41.6:	D8 1,	Length 7218:	
	Best Local Similarity	2.9%:	Pred No. 0.081:			
	Matches 11:	Conservative 210:	Mismatches 159:	Indels 0:	Gaps 0:	
OY	14	aataacaattgacaaagggtgaagaatcctaacaaggtatgaagccagtgtcca	73			
		: : : : : : : : : : :				
Db	1451	ATAGAGAATTGTAACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR	1392			
OY	74	ggctgcattcggttcaacgaactgcccacagacttgcatygaactygaaatygtata	133			
		: : : : : : : : : : : : : : : : :				
Db	1331	RRR	1332			
OY	134	attacaagaagaacaggyagacttaigtcaagaagagacgaigtlytcagcgaca	193			
		: : : : : : : : : : : : : : : : :				
Db	1331	RRR	1272			
OY	194	gcaccttagttcttcataattatatatgatalgtatcttatatatatatattatat	253			
		: : : : : : : : : : : : : : : : :				
Db	1271	RRR	1212			
OY	254	attaccatccgaagtaccocagtcactcgtaccattccocagvgagacatggtctcc	313			
		: : : : : : : : : : : : : : : : :				
Db	1211	RRR	1152			
OY	314	aagcgacacagaaaggttagcaggaagggcacgacgtgcacgtcggggctgt	373			
		: : : : : : : : : : : : : : : : :				
Db	1151	RRR	1092			
OY	374	gctcacagaagctgcagag	393			
		: : : : : : : : : : :				
Db	1091	RRR	1072			

RESULT 2  
 PCT-US93-06251-63  
 : Sequence 63, Application PC/TUS9306251  
 : GENERAL INFORMATION:  
 : APPLICANT: Wickstrom, Eric and Rife, Jason P.  
 : TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing  
 : TITLE OF INVENTION: Stereospecific Alkylphosphonates and Arylphosphonates  
 : NUMBER OF SEQUENCES: 93  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
 : STREET: 400 Garden City Plaza  
 : CITY: Garden City  
 : STATE: NY  
 : COUNTRY: USA  
 : ZIP: 11530  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: Patentin Release #1.0, Version #1.25  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: PCT/US93/06251  
 : FILING DATE: 19930630  
 : CLASSIFICATION:  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: Digilio, Frank S.  
 : REGISTRATION NUMBER: 31,346  
 : REFERENCE/DOCKET NUMBER: 8586  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: 516-742-4343  
 : TELEFAX: 516-742-4366  
 : TELEX: 230 901 SANS UR  
 : INFORMATION FOR SEQ ID NO: 63:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 2875 base pairs  
 : TYPE: nucleic acid  
 : STRANDEDNESS: double  
 : TOPOLOGY: linear  
 : MOLECULE TYPE: DNA (genomic)  
 : PCT-US93-06251-63

```

Query Match 7.0%; Score 38.8; DB 5; Length 2875;
Best Local Similarity 47.5%; Pred. No. 0.31;
Matches 115; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

OY 57 attgagccagtgltccaggctgcagctcagttcacagaactgltccaggaagctgcacag 116
Db 1671 AATCGACTAAATGGCATTGCTATCTTCATGTGACCCAAAAAATCTTAGGAAACCCCTGAAG 1730

OY 117 aactgaaactgltacataatcaagaagaacaaggaggaacttagtcgacagagaagacg 176
Db 1731 AAAAGAAACTGAGACTAATTTTATAGAGAAATGAGAGAGAGTTTGCATGAGAACTGAGAA 1790

OY 177 agtctggaaggcggaacagcacccttaagtccttcacattatataatgcatatgtatc 236
Db 1791 AGAGAAATCAATCCCAAGCAAGTTTATTATACAAATTTTAAATATGGTAAAGGTTTACT 1850

OY 237 tatatatattatataatttataatccagatccagatccagatccagatccatccatccag 296
Db 1851 TTGAATGATAAATTTCTATAGTTTAGTGTTAGTGAATCTTAACCTGTAATGTTGAACATTTGCTCATG 1910

OY 297 ga 298
Db 1911 CA 1912

RESULT 3
US-09-443-501A-2/c
; Sequence 2, Application US/09443501A
; Patent No. 6303342
; GENERAL INFORMATION:
; APPLICANT: Kosan Biosciences, Inc.
; APPLICANT: Julien, Bryan

```

[illegible]











```

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/483,554
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/348,824
FILING DATE: 29-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/308,104
FILING DATE: 16-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/300,266
FILING DATE: 02-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/289,221
FILING DATE: 12-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ihnen, Jeffrey L.
REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: 24884-109347
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-962-4810
TELEFAX: 202-962-8300
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 6769 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-08-850-727-20

Query Match          6.3%   Score 35.2; DB 4; Length 6769;
Best Local Similarity 58.7%; Pred.No.4.4;
Matches 61; Conservative 0; Mismatches 43; Indels 0; Gaps 0

QY    204 tcttcacattatagtatgatatgtatttccatataatatacatattttaccac 263
      ||||| ||||| ||||| | ||||| ||||| ||||| ||||| |||||
Db    1206 TATGTTAATAATATATCATCATATCATATATATATATATATATATATATATA 1147
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY    264 cagtatcccagtcactcgtlaccatttccccaggsgaacatgggt 307
      ||||| | |||| | |||| | |||| | |||| | |||| | ||||
Db    1146 TATATATATATATATGTAAATCCACGACACTTTCGGATATATGTGT 1103

RESULT 12
; Sequence 20, Application PC/TUS9510202
; GENERAL INFORMATION:
; APPLICANT: Shattuck-Eidens, Donna M.
; APPLICANT: Sismard, Jacques
; APPLICANT: Emi, Mitsuru
; APPLICANT: Nakamura, Yusuke
; APPLICANT: Durocher, Francine
; TITLE OF INVENTION: In Vivo Mutations and Polymorphisms
; TITLE OF INVENTION: in the 17q-Linked Breast and Ovarian Cancer
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESS: Venable, Baetjer, Howard & Civiletti, LLP
; STREET: 1201 New York Avenue, N.W., Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/10202
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/409,305
FILING DATE: 24-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/348,824
FILING DATE: 29-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08-308,104
FILING DATE: 16-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/300,266
FILING DATE: 02-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/289,221
FILING DATE: 12-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ihnen, Jeffrey L.
REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: 24884-109347
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-962-4810
TELEFAX: 202-962-8300
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 6769 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
PCT-US95-10202-20

Query Match          6.3%, Score 35.2; DB 5; Length 6769;
Best Local Similarity 58.7%; Pred. No. 4.4;
Matches   61; Conservative    0; Mismatches 45; Indels    0; Gaps    0;

QY      204 tcttcacattatatacgttatagttatttctatatatatatttaccac 263
        ||| ||||| ||||| ||||| | ||||||||| ||||| |||
DB      1206 TATGTGTAATATAATVTCACATATATCATCATATATATATATATATAA 1147
        ||| ||||| ||||| ||||| | ||||||||| ||||| |||

QY      264 caggatcccgactcatctgtaccattcccaggagacatg 307
        ||| ||||| ||||| ||||| | ||||| ||| ||
DB      1146 TATATATATATATATGTAAATCCACAGACTTTGGCATATATGTGT 1103
        ||| ||||| ||||| ||||| | ||||| ||| ||

RESULT 13
PCT-US95-10203-20/c
Sequence 20, Application PC/TUS9510203
GENERAL INFORMATION:
APPLICANT: Skolnick, Mark H.
APPLICANT: Goldgar, David E.
APPLICANT: MIKI, Yoshio
APPLICANT: Swenson, Jeff
APPLICANT: Kamb, Alexander
APPLICANT: Harshman, Keith D.
APPLICANT: Shattuck-Eidens, Donna M.
APPLICANT: Tavtigian, Sean V.
APPLICANT: Wiseman, Roger W.
APPLICANT: Futreal, P. Andrew
TITLE OF INVENTION: 17q-Linked Breast and Ovarian Cancer

```

```

? TITLE OF INVENTION: Susceptibility Gene
? NUMBER OF SEQUENCES: 85
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP
? STREET: 1201 New York Avenue, N.W., Suite 1000
? CITY: Washington
? STATE: DC
? COUNTRY: USA
? ZIP: 20005
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: PCT/US95/10203
? FILING DATE:
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US
? FILING DATE: 07-JUN-1995
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/409,305
? FILING DATE: 24-MAR-1995
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/348,824
? FILING DATE: 29-NOV-1994
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08-308,104
? FILING DATE: 16-SEP-1994
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/300,266
? FILING DATE: 02-SEP-1994
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/289,221
? FILING DATE: 12-AUG-1994
? ATTORNEY/AGENT INFORMATION:
? NAME: Ihnen, Jeffrey L.
? REGISTRATION NUMBER: 28,957
? REFERENCE/DOCKET NUMBER: 24884-109347
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 202-962-4810
? TELEFAX: 202-962-8300
? INFORMATION FOR SEQ ID NO: 20:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 6769 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: linear
? MOLECULE TYPE: DNA (genomic)
? HYPOTHETICAL: NO
? ANTI-SENSE: NO
? ORIGINAL SOURCE:
? ORGANISM: Homo sapiens
? PCT-US95-10203-20

Query Match 6.3%; Score 35.2; DB 5; Length 6769;
Best local Similarity 58.7%; Pred. No. 4.4;
Matches 61; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 204 tcttcataattatataatgatatattctatataattatattacac 263
      | | | | | | | | | | | | | | | | | | | | | | | |
DB 1206 TATGTATATATATATATATATATATATATATATATATATATATATATATA 1147

QY 264 caggatccagcatcgtacacattccaggagagacatgggt 307
      | | | | | | | | | | | | | | | | | | | | | |
DB 1146 TATATATATATATATATATATATATATATATATATATATATATATATAT 1103

RESULT 14
PCT-US95-10220-20/C
; Sequence 20, Application PC/TUS9510220
```

```

? GENERAL INFORMATION:
? APPLICANT: Skolnick, Mark H.
? APPLICANT: Goldgar, David E.
? APPLICANT: Miki, Yoshio
? APPLICANT: Swenson, Jeff
? APPLICANT: Kamb, Alexander
? APPLICANT: Harshman, Keith D.
? APPLICANT: Shattuck-Eidens, Donna M.
? APPLICANT: Tavtigian, Sean V.
? APPLICANT: Wiseman, Roger W.
? APPLICANT: Futreal, P. Andrew
? TITLE OF INVENTION: Method for Diagnosing a
? TITLE OF INVENTION: Predisposition for Breast and Ovarian Cancer
? NUMBER OF SEQUENCES: 85
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP
? STREET: 1201 New York Avenue, N.W., Suite 1000
? CITY: Washington
? STATE: DC
? COUNTRY: USA
? ZIP: 20005
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: PCT/US95/10220
? FILING DATE:
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US
? FILING DATE: 07-JUN-1995
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/409,305
? FILING DATE: 24-MAR-1995
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/348,824
? FILING DATE: 29-NOV-1994
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08-308,104
? FILING DATE: 16-SEP-1994
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/300,266
? FILING DATE: 02-SEP-1994
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/289,221
? FILING DATE: 12-AUG-1994
? ATTORNEY/AGENT INFORMATION:
? NAME: Ihnen, Jeffrey L.
? REGISTRATION NUMBER: 28,957
? REFERENCE/DOCKET NUMBER: 24884-109347
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 202-962-4810
? TELEFAX: 202-962-8300
? INFORMATION FOR SEQ ID NO: 20:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 6769 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: linear
? MOLECULE TYPE: DNA (genomic)
? HYPOTHETICAL: NO
? ANTI-SENSE: NO
? ORIGINAL SOURCE:
? ORGANISM: Homo sapiens
? PCT-US95-10220-20

Query Match 6.3%; Score 35.2; DB 5; Length 6769;
Best local Similarity 58.7%; Pred. No. 4.4;
Matches 61; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
```



**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 9, 2002, 15:22:36 ; Search time 852.09 Seconds

(without alignments)  
1118,294 Million cell updates/sec

Title: US-09-853-544-1

Perfect score: 1 tttttttttttaataac.....agtgtgcagcgtgcacaca 555

Sequence: 1 tttttttttttaataac.....agtgtgcagcgtgcacaca 555

Scoring table: IDENTITY\_NUC

Searched: 1736436 segs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 1008  
Listing first 45 summaries

Database : N\_Geneseq.032802.\*

```
1: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2002.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
C 1	531.6	95.8	1562	22 ABA15366 Human nervous syst
C 2	531.6	95.8	1562	22 ABA15367 Human nervous syst
C 3	531.6	95.8	1562	22 AAS27703 DNA encoding novel
C 4	531.6	95.8	1562	22 AAS27706 DNA encoding novel
C 5	531.6	95.8	4169	22 AAH77992 Nucleotide sequenc
C 6	528.4	95.2	1002	22 AAH99684 Human protein enco
C 7	77.6	14.0	272	22 ABA11460 Human nervous syst
C 8	63	11.4	5241	22 AAH77993 Nucleotide sequenc
C 9	43.2	7.8	557	22 AAH10423 Human CDNA clone (

C 10	43.2	7.8	2763	22 AAH17538
C 11	40.2	7.2	1599	23 ABL08153
C 12	40.2	7.2	6229	23 ABL08152
C 13	39.6	7.1	2914	22 AAK70730
C 14	39	7.0	15548	22 ABL34155
C 15	38.8	7.0	2888	11 AAO03743
C 16	38.6	7.0	6754	23 AAS61304
C 17	38	6.8	14061	23 ABL03394
C 18	38	6.8	16502	22 ABL32727
C 19	37.4	6.7	1140	22 ABL71781
C 20	37.4	6.7	1716	22 AA171755
C 21	37.4	6.7	2217	22 ABA82698
C 22	37.4	6.7	2320	21 AAF21879
C 23	37.4	6.7	3675	23 ABL14664
C 24	37.4	6.7	4320	23 ABL04688
C 25	37.4	6.7	21311	23 ABL04750
C 26	37.2	6.7	5379	24 ABL33677
C 27	37.2	6.7	5379	24 ABL34577
C 28	37.2	6.7	138169	21 ABA34791
C 29	37.2	6.7	14189	21 AAF20913
C 30	37.2	6.7	14189	21 AAF21127
C 31	37.2	6.7	14189	21 AAF21152
C 32	37.2	6.7	141589	21 AAA35005
C 33	37.2	6.7	141589	21 AAA35030
C 34	37.2	6.7	146981	21 AAF21442
C 35	37.2	6.7	209273	21 AAF21437
C 36	37	6.7	1341	21 AAC76115
C 37	36.8	6.6	71989	21 AAA29349
C 38	36.6	6.6	1877	22 AAS02381
C 39	36.6	6.6	5379	24 ABL33676
C 40	36.6	6.6	5379	24 ABL34576
C 41	36.6	6.6	32190	22 AAS29937
C 42	36.6	6.6	113515	24 ABL34174
C 43	36.4	6.6	10553	22 ABA16395
C 44	36.4	6.6	10553	22 AAK69113
C 45	36.2	6.5	3730	22 ABA19745

#### ALIGNMENTS

RESULT 1	ABAI5366/c	standard; DNA; 1562 BP.
ID	ABAI5366;	
XX	ABAI5366;	
AC	ABAI5366;	
XX	ABAI5366;	
DT	23-JUN-2002	(first entry)
XX	23-JUN-2002	
DE	Human nervous system related polynucleotide SEQ ID NO 7697.	
XX	Human; noctropic; neuroprotective; cytosolic; dermatological; viral; immunosuppressive; anti-inflammatory; anti-HIV; antibacterial; vulnary;	
KW	antiparkinsonian; antiscaling; antianaemic; antirheumatic; cancer;	
KW	antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;	
KW	antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;	
KW	antiparasitic; cardiac; immune disorder; cardiovascular disorder;	
KW	neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.	
XX		
OS	Homo sapiens.	
PN	WO200159063-A2.	
XX	16-AUG-2001.	
PD	16-AUG-2001.	
XX	17-JAN-2001; 2001WO-US01334.	
PF	17-JAN-2001; 2001WO-US01334.	
XX	31-JAN-2000; 2000US-0179065.	
PR	04-FEB-2000; 2000US-0180628.	
PR	24-FEB-2000; 2000US-0184664.	
PR	02-MAR-2000; 2000US-0186350.	
PR	16-MAR-2000; 2000US-0189874.	
PR	17-MAR-2000; 2000US-0190076.	

Human CDNA sequenc  
Drosophila melanog  
Drosophila melanog  
Human immune/haema  
Human immune/haema  
Human SKI related  
Human gene regulat  
Drosophila melanog  
Human immune syste  
Human zyxine fragm  
Human zyxine codin  
zyxin gene SEQ ID  
Human breast and o  
Drosophila melanog  
Drosophila melanog  
Drosophila melanog  
Human immune syste  
Human metastasis a  
Human adenosine re  
Human ELAM-1 polyn  
Human ELAM-1 polyn  
Human low adenosin  
Human low adenosin  
Human adenosine re  
Human adenosine re  
Human ELAM-1 polyn  
Human Factor-relat  
Human OREF ORF1670  
Sorangiium cellulosa  
Brassica napus DNA  
Human immune syste  
Human metastasis a  
Human lung antigen  
Human immune syste  
Human nervous syst  
Human immune/haema  
Human nervous syst

PR	18-MAR-2000	2000US-0198123
PR	19-MAR-2000	2000US-0205515
PR	07-JUN-2000	2000US-0209467
PR	28-JUN-2000	2000US-0211486
PR	30-JUN-2000	2000US-0215135
PR	07-JUL-2000	2000US-0216647
PR	07-JUL-2000	2000US-0216680
PR	11-JUL-2000	2000US-0217487
PR	11-JUL-2000	2000US-0217496
PR	14-JUL-2000	2000US-0218290
PR	26-JUL-2000	2000US-0220964
PR	26-JUL-2000	2000US-0220963
PR	14-AUG-2000	2000US-0224518
PR	14-AUG-2000	2000US-0224519
PR	14-AUG-2000	2000US-0225213
PR	14-AUG-2000	2000US-0225214
PR	14-AUG-2000	2000US-0225266
PR	14-AUG-2000	2000US-0225267
PR	14-AUG-2000	2000US-0225268
PR	14-AUG-2000	2000US-0225270
PR	14-AUG-2000	2000US-0225547
PR	14-AUG-2000	2000US-0225577
PR	14-AUG-2000	2000US-0225752
PR	14-AUG-2000	2000US-0225758
PR	18-AUG-2000	2000US-0225759
PR	18-AUG-2000	2000US-0226279
PR	22-AUG-2000	2000US-0226681
PR	22-AUG-2000	2000US-0226688
PR	22-AUG-2000	2000US-0227182
PR	23-AUG-2000	2000US-0227070
PR	30-SEP-2000	2000US-0229824
PR	01-SEP-2000	2000US-0229827
PR	01-SEP-2000	2000US-0229343
PR	01-SEP-2000	2000US-0229344
PR	01-SEP-2000	2000US-0229345
PR	05-SEP-2000	2000US-0229509
PR	05-SEP-2000	2000US-0229513
PR	06-SEP-2000	2000US-0230437
PR	06-SEP-2000	2000US-0230438
PR	08-SEP-2000	2000US-0231142
PR	08-SEP-2000	2000US-0231143
PR	08-SEP-2000	2000US-0231144
PR	08-SEP-2000	2000US-0231145
PR	08-SEP-2000	2000US-0232080
PR	08-SEP-2000	2000US-0232081
PR	12-SEP-2000	2000US-0231968
PR	14-SEP-2000	2000US-0232397
PR	14-SEP-2000	2000US-0232398
PR	14-SEP-2000	2000US-0232399
PR	14-SEP-2000	2000US-0232401
PR	14-SEP-2000	2000US-0232401
PR	14-SEP-2000	2000US-0233063
PR	14-SEP-2000	2000US-0233064
PR	21-SEP-2000	2000US-0234223
PR	21-SEP-2000	2000US-0234224
PR	25-SEP-2000	2000US-0234997
PR	25-SEP-2000	2000US-0234998
PR	26-SEP-2000	2000US-0235684
PR	27-SEP-2000	2000US-0235684
PR	27-SEP-2000	2000US-0235636
PR	29-SEP-2000	2000US-0236127
PR	29-SEP-2000	2000US-0236127
PR	29-SEP-2000	2000US-0236167
PR	29-SEP-2000	2000US-0236168
PR	29-SEP-2000	2000US-0236570
PR	29-SEP-2000	2000US-0236570
PR	02-OCT-2000	2000US-0237038
PR	02-OCT-2000	2000US-0237039
PR	02-OCT-2000	2000US-0237040
PR	13-OCT-2000	2000US-0239935
PR	13-OCT-2000	2000US-0239937
PR	13-OCT-2000	2000US-0239937

PR		20-OCT-2000;	2000US-0240960.
PR		20-OCT-2000;	2000US-0241785.
PR		20-OCT-2000;	2000US-0241786.
PR		20-OCT-2000;	2000US-0241787.
PR		20-OCT-2000;	2000US-0241808.
PR		20-OCT-2000;	2000US-0241809.
PR		20-OCT-2000;	2000US-0241826.
PR		20-OCT-2000;	2000US-0242221.
PR		01-NOV-2000;	2000US-0244617.
PR		08-NOV-2000;	2000US-0246475.
PR		08-NOV-2000;	2000US-0246476.
PR		08-NOV-2000;	2000US-0246477.
PR		08-NOV-2000;	2000US-0246478.
PR		08-NOV-2000;	2000US-0246523.
PR		08-NOV-2000;	2000US-0246524.
PR		08-NOV-2000;	2000US-0246525.
PR		08-NOV-2000;	2000US-0246526.
PR		08-NOV-2000;	2000US-0246527.
PR		08-NOV-2000;	2000US-0246528.
PR		08-NOV-2000;	2000US-0246532.
PR		08-NOV-2000;	2000US-0246609.
PR		08-NOV-2000;	2000US-0246610.
PR		08-NOV-2000;	2000US-0246611.
PR		08-NOV-2000;	2000US-0246613.
PR		17-NOV-2000;	2000US-0249207.
PR		17-NOV-2000;	2000US-0249208.
PR		17-NOV-2000;	2000US-0249215.
PR		17-NOV-2000;	2000US-0249216.
PR		17-NOV-2000;	2000US-0249217.
PR		17-NOV-2000;	2000US-0249218.
PR		17-NOV-2000;	2000US-0249244.
PR		17-NOV-2000;	2000US-0249245.
PR		17-NOV-2000;	2000US-0249264.
PR		17-NOV-2000;	2000US-0249265.
PR		17-NOV-2000;	2000US-0249297.
PR		17-NOV-2000;	2000US-0249299.
PR		17-NOV-2000;	2000US-0249300.
PR		01-DEC-2000;	2000US-0250391.
PR		01-DEC-2000;	2000US-0251160.
PR		05-DEC-2000;	2000US-0251030.
PR		05-DEC-2000;	2000US-0251988.
PR		05-DEC-2000;	2000US-0251989.
PR		06-DEC-2000;	2000US-0256179.
PR		06-DEC-2000;	2000US-0251479.
PR		08-DEC-2000;	2000US-0251856.
PR		08-DEC-2000;	2000US-0251868.
PR		08-DEC-2000;	2000US-0251869.
PR		08-DEC-2000;	2000US-0251899.
PR		11-DEC-2000;	2000US-0251990.
PR		11-DEC-2000;	2000US-0254097.
PR		05-JAN-2001;	2001US-0259678.
PA	(HUMA-)	HUMAN GENOME SCI INC.	
PI	Rosen CA,	Barash SC,	Ruben SM;
XX	WPI;	2001-541565/60.	
DR			
XX			
XX	Nucleic acids encoding	3224 human nervous system antigen polypeptides	
PT	useful for preventing,	diagnosing and/or treating nervous system	
PT	cancers and metastases -		
XX			
PS	Disclosure:	SEQ ID NO 7697; 1701pp + Sequence Listing; English.	
CC	The invention relates to novel genes (ABA11004-ABA21534) and proteins		
CC	(ABBI4678-ABBI8001) useful for preventing, treating or ameliorating		
CC	medical conditions e.g. by protein or gene therapy. The genes are		

CC isolated from a range of human tissues disclosed in the specification.  
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful  
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast  
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone  
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
CC haemolytic anemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;  
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
CC and parasitic infections.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://wipo.int/pub/published_pct_sequences).

CC  
XX  
SQ Sequence 1562 BP; 311 A; 529 C; 423 G; 299 T; 0 other;

Query Match 95.8%; Score 531.6; DB 22; Length 1562;  
Best Local Similarity 98.9%; Pred. No. 4,9e-133;  
Matches 545; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1 tttttttcttaataaattgacaaaaggtgaagaatcctaacaaggtatg 60  
DB 1514 tttttttcttaataaattgacaaaaggtgaagaatcctaacaaggtatg 1455  
QY 61 aggcacagtgccagcgtcattcagtcacagacgtccctcagacgttcagtgtaact 140  
DB 1454 aggcacagtgccagcgtcattcagtcacagacgtccctcagacgttcagtgtaact 1395  
QY 121 ggaatgtgtataaattacagaanaaaacaggaggaacttagtcagagagagagagtg 180  
DB 1394 ggaatgtgtataaattacagaanaaaacaggaggaacttagtcagagagagagagtg 1335  
QY 181 tggacggagcaacagcactcctttagcttattatattatagatgatatgattttata 240  
DB 1334 tggacggagcaacagcactcctttagcttattatattatagatgatatgattttata 1275  
QY 241 tatatatattatattatattatattatattatattatattatattatattatattat 300  
DB 1274 tatatatattatattatattatattatattatattatattatattatattatattat 1215  
QY 301 catgggtgcttccaaagcgagagaaaggttagcagagagagagagcgagtgctc 360  
DB 1214 catgggtgcttccaaagcgagagaaaggttagcagagagagagagcgagtgctc 1155  
QY 361 aggcgtgggcttggtcacagaagctgcagagagcttcacagagctgttaagagggcccccgg 420  
DB 1154 aggcgtgggcttggtcacagaagctgcagagagcttcacagagctgttaagagggcccccgg 1095  
QY 421 ctccgagaagccaggtact-gagncaaagccagctcctcagctcagcccccgtgctg 479  
DB 1094 ctccgagaagccaggtact-gagncaaagccagctcctcagctcagcccccgtgctg 1035  
QY 480 atccaccgcttcccgcaacttcacatcatcagagggcccgcttcattgcatcagtg 539  
DB 1034 atccaccgcttcccgcaacttcacatcatcagagggcccgcttcattgcatcagtg 975  
QY 540 gtgcagcgtgc 550  
DB 974 gtgcagcgtgc 964

RESULT 2  
ID ABA15367/C  
ABAI5367 standard; DNA; 1562 BP.

XX ABA15367;  
XX  
DT 23-JAN-2002 (first entry)  
XX  
DE Human nervous system related polynucleotide SEQ ID NO 7698.

XX Human; nootropic; neuroprotective; cytosolic; dermatological; virucide;  
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;  
KW antiparkinsonian; antischizoid; antianaemic; antiarthritic; cancer;  
KW antihemetic; hepatotropic; cerebroprotective; antiinflammatory;  
KW antiallergic; antidiabetic; antilucer; anticonvulsant; antifungal;  
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.  
XX  
XX Homo sapiens.  
XX  
XX WO200159063-A2.  
PD 16-AUG-2001.  
XX  
XX  
PF 17-JAN-2001; 2001WO-US01334.  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226379.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226688.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 06-SEP-2000; 2000US-0231142.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.

PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 13-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 20-OCT-2000; 2000US-0242221.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250391.  
PR 01-DEC-2000; 2000US-0251160.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.

PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
PI Rosen CA, Barash SC, Ruben SM.  
XX  
XX WPI; 2001-541565/60.  
DR  
XX  
XX  
PT Nucleic acids encoding 3224 human nervous system antigen polypeptides,  
XX useful for preventing, diagnosing and/or treating nervous system  
XX cancers and metastases -  
PS  
XX Disclosure; SEQ ID NO 7698; 1701bp + Sequence Listing; English.  
XX  
XX  
XX The invention relates to novel genes (ABAI1004-ABA21534) and proteins  
CC (ABAI4678-ABAI8001) useful for preventing, treating or ameliorating  
CC medical conditions e.g. by protein or gene therapy. The genes are  
CC isolated from a range of human tissues disclosed in the specification.  
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful  
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast  
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone  
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;  
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
CC and parasitic infections.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pat\_sequences.  
XX  
XX  
XX Sequence 1562 BP; 312 A; 528 C; 423 G; 299 T; 0 other;  
SQ  
Query Match 95.8%; Score 531.6; DB 22; Length 1562;  
Best Local Similarity 98.9%; Pred. No. 4,9e-133;  
Matches 545; Conservative 0; Mismatches 5; Indels 1; Gaps 1;  
QY 1 ttttttttcttaataacaatttgacaaagggtgaagaagaatcctaacaagatgac 60  
DB 1514 ttttttttcttaataatacatttgacanaagggtgacanaaatcctaanaagatgac 1455  
QY 61 aggcacagtcacagctgcattcagtcacagaactgtccctcagagcgttcagtggaact 120  
DB 1454 aggcacagtcgacagctgcattcagtcacagaactgtccctcagagcgttcagtggaact 1395  
QY 121 ggaatggtatataataacagaagaacagagggagacttagtcagaaggaggaagatg 180  
DB 1394 ggaatggtatataataatacagaagaacagagggagacttagtcagaaggaggaagatg 1335  
QY 181 tggacgggcaacagacatcccttagcttctcatattatataatgatatgtaattctata 240  
DB 1334 tggacgggcaacagacatcccttagcttctcatattatataatgatatgtaattctata 1275  
QY 241 tatatatatatattatcatccagtgatccagtcacatcgtaccattcccaaggagaga 300  
DB 1274 tatatatatatattatatttaccatccagtgatccagtcacatcgtaccattcccaaggagaga 1215  
QY 301 catgggtgcttcccaaggcgagacagaagaagggttaaggcaggaaggagcagcaggtgc 360  
DB 1214 catgggtgcttcccaaggcgagacagaagaagggttaaggcaggaaggagcagcaggtgc 1155  
QY 361 aggcctgggcttggtccacagaagctcagaggaactccaacagatgtaagaagcccccggg 420  
DB 1154 aggcctgggcttggtccacagaagctcagaggaactccaacagatgtaagaagcccccggg 1095

QY 421 ctccgcagacgcccaggtact-gaguncaaaagccagtcctccagctccacgcccgcctg 479  
 |||||||  
 Db 1094 CTCGCAGACGCCGAGTACTGTGACAGCAAGCCAGTCTCCAGCTCAGCCCGGCTGG 1035  
 QY 480 atccacgcgctcttcgcgaacttcatacatcagggccgcgctcactgctgacatc 539  
 |||||||  
 Db 1034 ATCCACCGCTTCTCCGCAACTTCATCATCATCAGGCGCCGCTTCATGTCGATCAGTT 975  
 QY 540 gtgcagcgtgc 550  
 |||||||  
 Db 974 GTGCAGCCTGC 964

RESULT 3  
 ID AAS27703/c  
 AAS27703 standard; DNA; 1562 Bp.

XX AAS27703;  
 AC  
 XX  
 DT 07-NOV-2001 (first entry)  
 XX  
 DE DNA encoding novel signal transduction pathway protein, Seq ID 1363.  
 XX  
 KW Neuroprotective; cytoskeletal; dermatological; immunosuppressive; tumour;  
 KW antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;  
 KW immune system disorder; rheumatoid arthritis; inflammatory condition;  
 KW organ transplant rejection; infection; hepatitis C; blood disorder;  
 KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;  
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
 KW chromosomal abnormality; Down syndrome; Ischaemia; renal disorder;  
 KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;  
 KW reproductive system; gastrointestinal; liver disorder; AIDS; ds;  
 KW acquired immune deficiency syndrome.

XX Homo sapiens.  
 OS  
 EN W0200154733-A1.  
 XX  
 PD 02-AUG-2001.  
 XX  
 PF 17-JAN-2001; 2001WO-US01312.  
 XX  
 PR 31-JAN-2000; 2000US-0179065.  
 PR 04-FEB-2000; 2000US-0180628.  
 PR 24-FEB-2000; 2000US-0184664.  
 PR 02-MAR-2000; 2000US-0186350.  
 PR 16-MAR-2000; 2000US-0189874.  
 PR 17-MAR-2000; 2000US-0190076.  
 PR 18-APR-2000; 2000US-0198122.  
 PR 19-MAY-2000; 2000US-0205515.  
 PR 07-JUN-2000; 2000US-0209467.  
 PR 28-JUN-2000; 2000US-0214886.  
 PR 30-JUN-2000; 2000US-0215135.  
 PR 07-JUL-2000; 2000US-0216647.  
 PR 07-JUL-2000; 2000US-0216880.  
 PR 11-JUL-2000; 2000US-0217487.  
 PR 11-JUL-2000; 2000US-0217496.  
 PR 14-JUL-2000; 2000US-0218290.  
 PR 26-JUL-2000; 2000US-0220963.  
 PR 26-JUL-2000; 2000US-0220964.  
 PR 14-AUG-2000; 2000US-0224518.  
 PR 14-AUG-2000; 2000US-0224519.  
 PR 14-AUG-2000; 2000US-0225213.  
 PR 14-AUG-2000; 2000US-0225214.  
 PR 14-AUG-2000; 2000US-0225266.  
 PR 14-AUG-2000; 2000US-0225267.  
 PR 14-AUG-2000; 2000US-0225268.  
 PR 14-AUG-2000; 2000US-0225270.  
 PR 14-AUG-2000; 2000US-0225447.  
 PR 14-AUG-2000; 2000US-0225757.  
 PR 14-AUG-2000; 2000US-0225758.  
 PR 14-AUG-2000; 2000US-0225759.

PR 18-AUG-2000; 2000US-0226279.  
 PR 22-AUG-2000; 2000US-0226681.  
 PR 22-AUG-2000; 2000US-0226868.  
 PR 22-AUG-2000; 2000US-0227182.  
 PR 23-AUG-2000; 2000US-0227109.  
 PR 30-AUG-2000; 2000US-0228924.  
 PR 01-SEP-2000; 2000US-0229287.  
 PR 01-SEP-2000; 2000US-0229343.  
 PR 01-SEP-2000; 2000US-0229344.  
 PR 01-SEP-2000; 2000US-0229345.  
 PR 05-SEP-2000; 2000US-0229509.  
 PR 05-SEP-2000; 2000US-0229513.  
 PR 06-SEP-2000; 2000US-0230437.  
 PR 06-SEP-2000; 2000US-0230438.  
 PR 08-SEP-2000; 2000US-0231242.  
 PR 08-SEP-2000; 2000US-0231243.  
 PR 08-SEP-2000; 2000US-0231244.  
 PR 08-SEP-2000; 2000US-0231413.  
 PR 08-SEP-2000; 2000US-0231414.  
 PR 08-SEP-2000; 2000US-0232080.  
 PR 08-SEP-2000; 2000US-0232081.  
 PR 12-SEP-2000; 2000US-0233168.  
 PR 14-SEP-2000; 2000US-0233397.  
 PR 14-SEP-2000; 2000US-0233399.  
 PR 14-SEP-2000; 2000US-0233399.  
 PR 14-SEP-2000; 2000US-0233400.  
 PR 14-SEP-2000; 2000US-0233401.  
 PR 14-SEP-2000; 2000US-0233063.  
 PR 14-SEP-2000; 2000US-0233064.  
 PR 14-SEP-2000; 2000US-0233065.  
 PR 21-SEP-2000; 2000US-0234223.  
 PR 21-SEP-2000; 2000US-0234274.  
 PR 25-SEP-2000; 2000US-0234998.  
 PR 25-SEP-2000; 2000US-0234998.  
 PR 26-SEP-2000; 2000US-0235484.  
 PR 27-SEP-2000; 2000US-0235834.  
 PR 27-SEP-2000; 2000US-0235836.  
 PR 29-SEP-2000; 2000US-0236327.  
 PR 29-SEP-2000; 2000US-0236367.  
 PR 29-SEP-2000; 2000US-0236368.  
 PR 29-SEP-2000; 2000US-0236369.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 02-OCT-2000; 2000US-0236802.  
 PR 02-OCT-2000; 2000US-0237037.  
 PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 02-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-0239335.  
 PR 13-OCT-2000; 2000US-0239337.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 01-NOV-2000; 2000US-0241826.  
 PR 01-NOV-2000; 2000US-0244617.  
 PR 08-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.

PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 05-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
PI WPI; 2001-465460/50.  
XX  
XX Novel polypeptides useful for diagnosing, treating, preventing and/or  
PT prognosing disorders related to the proteins, including cancers, immune  
PT disorders and neuronal disorders -  
XX  
XX Claim 1: SEQ ID NO 1363; 880bp; English.  
XX  
XX The invention relates to novel isolated polypeptides (I), and  
CC polynucleotides (II), (I), (II) and the antibody to (I) are useful for  
CC diagnosing, preventing and treating diseases including immune system  
CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune  
CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ  
CC transplant rejections and graft versus host disease, infectious diseases  
CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and  
CC other blood-related disorders (sickle cell anaemia), myeloproliferative  
CC disorders, primary haematopoietic disorders, hyperproliferative  
CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative  
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal  
CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal  
CC disorders (e.g. glomerulonephritis), cardiovascular disorders  
CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in  
CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.  
CC Addison's disease), reproductive system disorders, gastrointestinal  
CC disorder (inflammatory disorders), liver disorders (cirrhosis),  
CC as stimulators of B-cell responsiveness to pathogens, activators of  
CC T-cells, to induce higher affinity antibodies, and as a means to induce  
CC tumour proliferation in pathologies e.g. acquired immune deficiency  
CC syndrome (AIDS). AAS26976-AAS27850 represent novel signal transduction  
CC pathway protein coding sequences and PCR primers of the invention.  
XX  
XX Sequence 1562 BP; 311 A; 529 C; 423 G; 299 T; 0 other:  
SQ

Query Match 95.8%; Score 531.6; DB 22; Length 1562;  
Best Local Similarity 98.9%; Pred. NO. 4.9e-133;

Matches 545; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1 tttttttcttaataacaatttgacaaagggtgaagaatccctaacaaggtatcg 60  
|||||  
Db 1514 TTTTTCCTTTTAATTAACATTTTACAAAAGGGAAGAAATCTTAACAAGATTGG 1455  
|||||  
QY 61 aggcacagtgccagcgtcattcagtcacagaactgtccccaagcgttcagtaact 120  
|||||  
Db 1454 AGGCCAGTGTCCAGGCTGCATTTCAGTTCACAGAACTGCTCCAGAGCGTTGATGAACT 1395  
|||||  
QY 121 ggaatgtgtaataatacaagaagaagggaagctagtgacagaaggagacagtg 180  
|||||  
Db 1394 GGAAATGTATATATTAATTAAGAAAGGGAAGGAGGAGGAGGAGGAGGAGGAGT 1335  
|||||  
QY 181 tgaacgggcaacagcaactccttagcttcattatataatgataatgataatgata 240  
|||||  
Db 1334 TGGACGGGCAACAGCATCTTACTCTTTATATATATGTAATGTAATTTTCTATA 1275  
|||||  
QY 241 tatataattatataattatcatccaggtatcccaagtcacgtgacattcccaaggaga 300  
|||||  
Db 1274 TATATATTTATATATTTTATCATCCAGGTATCCAGTCACTGTAACCATTTCCAGGAGA 1215  
|||||  
QY 301 catggtgtctccaaaggcagaagcagaaggttaggcagggaaggagcagcagtg 360  
|||||  
Db 1214 CATGGTGTCTCCAAAGCGAGACAGGAAGGTTAAGCGAGGAAGGAGGAGGAGGAGGAGT 1155  
|||||  
QY 361 aggcctggagcttggtccacagaagctgcagagagcttcacagctgtaagaaggcccg 420  
|||||  
Db 1154 AGGCTGGGGCTTGGCTCAGCAAGAGCTGCAAGAGCTTCAACGACTGTAAAGAGGCCCGGG 1095  
|||||  
QY 421 ctccgcagaagccaggtact-gagmcaaaagcagtcctccagctccagcggcgtcg 479  
|||||  
Db 1094 CTCGCGAGAGCGCAGGTACTGTCAGCAAGGCCAAGCTCTCACCTCCACCCCGCTGGG 1035  
|||||  
QY 480 atccaccgctctccgcgaacttcacatcatcaaggccgcttcattgcatcagtt 539  
|||||  
Db 1034 ATCCACCGCCTTCTCCGAAACTTCATCATCAAGGCGCCGCTTCATGCAATCAGTT 975  
|||||  
QY 540 gtccagcgtgc 550  
|||||  
Db 974 GTGCAGCGTGC 964  
|||||

RESULT 4  
AAS27706/c  
ID AAS27706 standard; DNA: 1562 BP.  
XX  
XX AAS27706;  
XX  
XX 07-NOV-2001 (first entry)  
XX  
XX DNA encoding novel signal transduction pathway protein, Seq ID 1366.  
XX  
XX Neutropenictive; cytostatic; dermatological; immunosuppressive; tumour;  
KW antineoplastic; anti-HIV; antibacterial; anti-inflammatory; cancer;  
KW immune system disorder; rheumatoid arthritis; inflammatory condition;  
KW organ transplant rejection; infection; hepatitis C; blood disorder;  
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;  
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;  
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;  
KW reproductive system; gastrointestinal; liver disorder; AIDS; ds;  
KW acquired immune deficiency syndrome.  
XX  
XX Homo sapiens.  
OS  
OS  
OS WO200154733-A1.  
PN  
XX  
XX 02-AUG-2001.  
PD  
XX  
XX 17-JAN-2001; 2001WO-US01312.  
PF  
XX  
XX 31-JAN-2000; 2000US-0179065.  
PR

PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0233397.  
PR 14-SEP-2000; 2000US-0233398.  
PR 14-SEP-2000; 2000US-0233399.  
PR 14-SEP-2000; 2000US-0233400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.

PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241285.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
  
(HUMA-) HUMAN GENOME SCI INC.  
PA  
XX  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-465460/50.  
XX  
XX  
XX Novel polypeptides useful for diagnosing, treating, preventing and/or  
PT prognosing disorders related to the proteins, including cancers, immune  
PT disorders and neuronal disorders  
XX

PS Claim 1: SEQ ID NO 1366; 880bp; English.

XX The invention relates to novel isolated polypeptides (I), and  
 CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for  
 CC diagnosing, preventing and treating diseases including immune system  
 CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune  
 CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ  
 CC transplant rejections and graft versus host disease, infectious diseases  
 CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and  
 CC other blood-related disorders (sickle cell anaemia), myeloproliferative  
 CC disorders, primary haematopoietic disorders, hyperproliferative  
 CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative  
 CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal  
 CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal  
 CC disorders (e.g. glomerulonephritis), cardiovascular disorders  
 CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in  
 CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.  
 CC Addison's disease), reproductive system disorders, gastrointestinal  
 CC disorder (inflammatory disorders), liver disorders (cirrhosis),  
 CC as stimulators of B-cell responsiveness to pathogens, activators of  
 CC T-cells, to induce higher affinity antibodies, and as a means to induce  
 CC tumour proliferation in pathologies e.g. acquired immune deficiency  
 CC syndrome (AIDS). AAS26976-AAS27850 represent novel signal transduction  
 CC pathway protein coding sequences and PCR primers of the invention.

XX Sequence 1562 BP; 312 A; 528 C; 423 G; 299 T; 0 other;

Query Match 95.8%; Score 531.6; DB 22; Length 1562;  
 Best Local Similarity 98.9%; Pred. No. 4.9e-133;  
 Matches 545; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1 ttttttttcttaataacaattgcacaaagggtgaagaatctctaacaagatattg 60  
 DB 1514 TTTTCTTTCTTAATAACAAATTGACAAAGGTTGAAGAAATCTAAACAGATTG 1455  
 QY 61 aagccagtggtccaggtgcatcttcacagacgtctccagagcgttgatgaact 120  
 DB 1454 AGGCCAGTGCTCCAGCTGATTCAGTTCAAGAACTGCTCCAGAGCTTGACGAACT 1395  
 QY 121 ggaagtggtatataatacaagaagaacagggagactagtcagaagaggaacagtg 180  
 DB 1394 GGAAGTGTATATATTACAGAAAGAAACAGAGGACCTTAGTCACAGAGACACAGTG 1335  
 QY 181 tggagggcacaacagcactccttagcttcaatattatataatggtatatttctata 240  
 DB 1334 TGCACGGGCAACGACATCCTTACTCTTTATATATATGCTATATATTTCTATA 1275  
 QY 241 tatatatatatattatcatccaggtatcccaatcattcaccatttcccaaggaga 300  
 DB 1274 TATATATTTATATTTTACATCCAGGTATCCAGTCAATCTGACCAATTTCCACAGGAGA 1215  
 QY 301 catgggtgcttccaagcgagacagaagaagggttaaggcagggagaggagcagcagtg 360  
 DB 1214 CAGGGTGGTCCCAAGGCGAGACAGAAAGGTTAGGACAGGGAAGGCGACGACGCTGC 1155  
 QY 361 aggtctgggtcttggtccaagaagctgcaggagcttcagcagctgttaagaaggcccg 420  
 DB 1154 AGGCTGGGGCTTGCTCCACAGAACTGCAGAGACTTCAGAGATCTTAAAGAGGGCCCGG 1095  
 QY 421 ctccggaagcgccaggtact-gagncaaagccagctcctcagctcagcccccgttg 479  
 DB 1094 CTCGCGAGAGCGGAGTACTGGCAGCAAAAGCCAGTCTCCAGCTCCACGCGCCGCTGG 1035  
 QY 480 atccacgccttcctccgaactcatcatcagggcccgcttcattatgcatcagtt 539  
 DB 1034 ATCCACGCGCTTCTCCGCAAACTTCATCATCATCAGGCGCCGCTTATGTCATCCAGTT 975  
 QY 540 gtgcagcagtcgc 550  
 DB 974 GTGCACGCTGC 964

RESULT 5

AAH77992/C  
 ID AAH77992 standard; DNA; 4169 BP.  
 XX  
 AC AAH77992;  
 XX  
 DT 13-NOV-2001 (first entry)  
 XX  
 DE Nucleotide sequence of human protein kinase SGK223.

XX Human; protein kinase; cancer; immune disease; cardiovascular disease;  
 KW brain disease; neuronal disease; Alzheimer's disease; chromosome 8;  
 KW Parkinson's disease; multiple sclerosis; metabolic disorder;  
 KW peripheral nervous system disease; amyotrophic lateral sclerosis;  
 KW infection; ocular disease; migraine; pain; sexual dysfunction;  
 KW mood disorder; attention disorder; cognition disorder; hypotension;  
 KW hypertension; psychotic disorder; dyskinesia; transplant rejection; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..3735 /tag="a

FT /product="protein kinase"

PN WO200166594-A2.

XX 13-SEP-2001.

PD 02-MAR-2001; 2001WO-US06638.

XX 06-MAR-2000; 2000US-0187150.

PR 29-MAR-2000; 2000US-0193404.

PR 13-NOV-2000; 2000US-0247013.

PA (SUGEN) SUGEN INC.

PI Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;

XX WPI; 2001-536777/59.

DR P-PSDB; AAG67393.

XX Nucleic acids capable of encoding human polypeptides having a kinase or

PT kinase-like activity; useful for diagnosing a disease selected from

PT cancers, cardiovascular disease and neuronal-associated diseases (e.g.

PT Alzheimer's disease) -

XX Example 1; Fig 1B-C; 201pp; English.

PS The present sequence encodes a human protein kinase. The

CC gene is located at chromosomal position 8p22-p23. The kinase

CC polypeptides are useful for diagnosing a disease or disorder

CC selected from cancers (e.g. cancers of tissues and cancers of

CC hematopoietic origin), immune-related diseases and disorders,

CC cardiovascular disease, brain or neuronal-associated diseases (e.g.

CC Alzheimer's disease, Parkinson's disease, multiple sclerosis),

CC metabolic disorders, peripheral nervous system diseases, amyotrophic

CC lateral sclerosis, viral infections, infections caused by prions,

CC infections caused by bacteria, infections caused by fungi, ocular

CC diseases, migraines, pain, sexual dysfunction, mood disorders,

CC attention disorders, cognition disorders, hypotension, hypertension,

CC psychotic disorders, dyskinesias, and organ transplant rejection.

CC Kinase inhibitors are useful for treating diseases and disorders

CC described above.

XX Sequence 4169 BP; 871 A; 1398 C; 1167 G; 733 T; 0 other;

Query Match 95.8%; Score 531.6; DB 22; Length 4169;  
 Best Local Similarity 98.9%; Pred. No. 7.6e-133;  
 Matches 545; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1 ttttttttcttaataacaattgcacaaagggtgaagaatctctaacaagatattg 60  
 |||



Db	4110	TTTTTTTCTTTCTTAATTAACAATTTTGACAAAAGGGTGAAGAAATCTTAACAAAGGATTTG	4051
Qy	61	aggccagtgctcagggctgcatcagttcacagaaactgtccctcagagcgttcacatggaact	120
Db	4050	AGGCCAGTGTCOAGGCTGCATTTCAGTTCCAGAACTGTCTCCAGAGAGCTTCATGGAAC	3991
Qy	121	ggaatcgtctaatctacagaagaanaacagggaggaacttagtcagagagagagacagtg	180
Db	3990	GGAAATGTGTAAATTATACAGAAAGAAACAGGAGAGACTTAAGTCAGAGAGAGACGAGTG	3931
Qy	181	tggagggcacaacagcatccttagctcttcattatattatataatgatatatgtatttccata	240
Db	3930	TGGAGGGCAACAGCATCTTAGTCTTTCATATTAATGATATGATATTTTCTATA	3871
Qy	241	tatatattatattatattacatccaggtatccagatcctctgtatccattccagaggaga	300
Db	3870	TATATATTTATATATTTTACATCCAGGATATCCAGTACTCTGTACCATTTCCAGGGAGA	3811
Qy	301	catgggtgcttcacaaagcagacaggaagaggttagcagggagaaagggcagcgagtg	360
Db	3810	CATGGTGCTCTCCAGGCGAGACAGAGAAAGGGTTAAGCAGAGGAAGGGGACAGCGGATG	3751
Qy	361	aggtcgtgggtctgtgtctacacagaagctcagagagcttccagcgtgtaaagggcccg	420
Db	3750	AGGCTGGGGCTTGGCTCCAGAAAGTGTGAGAGGCTTTCAGCCACTGTAAAGGGGCCCGG	3691
Qy	421	ctccgcagaagccagaagtaact-gaggtcaaaagcaatctccctcagctccacggccgcctg	479
Db	3690	CTCCGCACACGCGCAAGTACTGGCAGCAAGCAAGCAGTCTCCAGTCCACGGCCCGCTGCG	3631
Qy	480	atccacgcgctcttcgcgaacttcattcatcatcatcagggccgcgcttcatgtcagtcag	539
Db	3630	ATCCACGCCGCTTCTCCGCAAACTCATCATCATATCATAGGGCCCGCTTCATGTGATCCAG	3571
Qy	540	gtcagcagtg	550
Db	3570	GTGACGCGTGC	3560
RESULT 6			
ID	AAH99684	standard; cDNA; 1002 BP.	
AC	AAH99684;		
XX			
DT	16-OCT-2001	(first entry)	
XX			
DE	Human protein encoding cDNA sequence SEQ ID NO:519.		
XX			
KW	Human; cancer; ulcer; HIV infection; human immunodeficiency virus;		
KW	antiflammatory; antirheumatic; antiarthritic; immunosuppressive;		
KW	antibacterial; endocrine; cardiac; central nervous system; virucide;		
KW	anti-HIV; fungicide; antimutagen; cardiovascular; antinaemic; anaemia;		
KW	antiaggregant; haemostatic; vulnery; antileuc; antileptic; eczema;		
KW	dermatological; antiallergic; antiaslumatic; antidiabetic; cytostatic;		
KW	neuroprotective; antidepressant; nootropic; antiparkinsonian; infection		
KW	immunostimulant; gene therapy; antisense therapy; vaccine; inflammation		
KW	antianaphalactic; rheumatoid arthritis; septic shock; pancreatitis;		
KW	cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;		
KW	genetic disease; haematopoietic disorder; platelet disorder; asthma;		
KW	thrombocytopaenia; osteoporosis; severe combined immunodeficiency;		
KW	allergic rhinitis; diabetes; multiple sclerosis; depression;		
KW	Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;		
KW	neurological disorder; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200153455-A2.		
XX			
PD	26-JUL-2001.		
XX			
PF	22-DEC-2000; 2000WO-US35017.		
XX			

PR 23-DEC-1999; 99US-0471275.  
PR 21-JAN-2000; 2000US-0488725.  
PR 25-APR-2000; 2000US-0552317.  
XX  
XX  
PA (HYSE-) HYSO INC.  
XX  
XX  
PI Tang YT, Liu C, Drmanac RT;  
XX  
XX WPI; 2001-457603/49.  
DR P-PSDB; AAM25743.  
DR  
XX  
XX Isolated human polynucleotides encoding polypeptides, useful for the  
PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection -  
PT  
XX  
XX Claim 1; Page 577; 1217pp; English.  
PS

AAH99166 to AAH99904 encode the human proteins given in AAM25225 to AAM25963. The proteins can have activities based on the tissues and cells they are expressed in, such as: antiinflammatory; antirheumatic; antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant; central nervous system; virocidic; anti-HIV; fungicide; antimutagen; cardiovascular; antilaemic; antiaggregant; haemostatic; vulnerary; antitumor; osteoprotic; dermatological; antiallergic; antiaslatic; antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic; antiparkinsonian; and immunostimulant. The proteins and polynucleotides encoding them can be used in gene therapy, antisense therapy and vaccine production. The proteins and polynucleotides are useful for screening for agonists or antagonists of a protein and for the treatment and diagnosis of disorders associated with the activity of a protein e.g. inflammation, rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction, neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal infections, autoimmunity, genetic diseases, haematopoietic disorders, anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers, osteoporosis, severe combined immunodeficiency, eczema, allergic rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression, Alzheimer's disease, Parkinson's disease, neurodegenerative and neurological disorders.

Query Match	95.2%	Score 528.4	DB 22	Length 1002
Best Local Similarity	98.5%	Pred. No. 2.9e-13		
Matches 543	Conservative 0	Mismatches 7	Indels 1	Gaps 1
QY 1	ttttttttctttaataacaatttgacaaaagggtgaagaactcctaacaagratg	60		
DB 40	ttttttttctttaataacaatttgacaaaagggtgaaaaaactctaacaagratg	99		
QY 61	aggccagtgctccaggctgcattcaattcacaagaactgctccaagacgtttgacatg	120		
DB 100	aggccagtgctccaggctgcattcaattcacaagaactgctccaagacgtttgacatg	159		
QY 121	ggaatggtataattatcaagaagaacacgggaagacttagtgcagagaggagacag	180		
DB 160	ggaatggtataattatcaagaagaacacgggaagacttagtgcagagaggagacag	219		
QY 181	tggacgggcaacagacatccctagctcttcataattatataatgatatattttctata	240		
DB 220	tggacgggcaacagacatccctagctcttcataattatataatgatatattttctata	279		
QY 241	tatatatttatattttaacatccaggtatcccaagtcacgttaaccattcccaaggaga	300		
DB 280	tatatatttatattttaacatccaggtatcccaagtcacgttaaccattcccaaggaga	339		
QY 301	catgggtgcttcccaaggcgacacagaagaagggttaggcagaggaaggcgacaggtgc	360		
DB 340	catgggtgcttcccaaggcgacacagaagaagggttaggcagaggaaggcgacaggtgc	399		
QY 361	aggctggggttggtctcacagaagctgcagagctcaagcagctgtaagaggccccgg	420		
DB 400	aggctggggttggtctcacagaagctgcagagctcaagcagctgtaagaggccccgg	459		

QY	421	ctccgcagaagccaggtact-gagncaaagccagtcctccagctccaaagcccccgtcg	479
Db	460		519
QY	480	atccaccgcctctccgcgaactcatcatcacaaggccgcctcatatgtcatccagtt	539
Db	520	atccaccgcctctccgcgaactcatcatcacaaggccgcctcatatgtcatccagtt	579
QY	540	gtgcagcgtgc 550	
Db	580	gtgcagcgtgc 590	
RESULT 7			
ABALL460/c			
ID	ABALL460 standard; cDNA; 272 BP.		
XX	ABALL460;		
AC			
XX	23-JAN-2002 (first entry)		
DT			
DE	Human nervous system related polynucleotide SEQ ID NO 467.		
XX			
KW	Human; nootropic; neuroprotective; cytosatic; dermatological; virucide;		
KW	immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;		
KW	antiparkinsonian; antischling; antianemic; antiarthritic; cancer;		
KW	antipneumatic; hepatotropic; cerebroprotective; antiinflammatory;		
KW	antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;		
KW	antiparasitic; cardiac; immune disorder; cardiovascular disorder;		
KW	neurological disease; infection; nephrotropic; gene therapy; vaccine; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200159063-A2.		
XX			
PD	16-AUG-2001.		
XX			
PF	17-JAN-2001; 2001WO-US01334.		
XX			
PR	31-JAN-2000;	2000US-0179065.	
PR	04-FEB-2000;	2000US-0180628.	
PR	24-FEB-2000;	2000US-0184664.	
PR	02-MAR-2000;	2000US-0186350.	
PR	16-MAR-2000;	2000US-0189674.	
PR	17-MAR-2000;	2000US-0190076.	
PR	18-APR-2000;	2000US-0198123.	
PR	19-MAY-2000;	2000US-0205515.	
PR	07-JUN-2000;	2000US-0209467.	
PR	28-JUN-2000;	2000US-0214886.	
PR	30-JUN-2000;	2000US-0215135.	
PR	07-JUL-2000;	2000US-0216647.	
PR	07-JUL-2000;	2000US-0216880.	
PR	11-JUL-2000;	2000US-0217487.	
PR	11-JUL-2000;	2000US-0217496.	
PR	14-JUL-2000;	2000US-0218290.	
PR	26-JUL-2000;	2000US-0220963.	
PR	26-JUL-2000;	2000US-0220964.	
PR	14-AUG-2000;	2000US-0224518.	
PR	14-AUG-2000;	2000US-0224519.	
PR	14-AUG-2000;	2000US-0224513.	
PR	14-AUG-2000;	2000US-0225214.	
PR	14-AUG-2000;	2000US-0225266.	
PR	14-AUG-2000;	2000US-0225267.	
PR	14-AUG-2000;	2000US-0225268.	
PR	14-AUG-2000;	2000US-0225270.	
PR	14-AUG-2000;	2000US-0225447.	
PR	14-AUG-2000;	2000US-0225757.	
PR	14-AUG-2000;	2000US-0225758.	
PR	14-AUG-2000;	2000US-0225759.	
PR	18-AUG-2000;	2000US-0226279.	
PR	22-AUG-2000;	2000US-0226681.	
PR	22-AUG-2000;	2000US-0226686.	
PR	22-AUG-2000;	2000US-0227182.	

PR	23-AUG-2000	2000US-0227009
PR	30-AGO-2000	2000US-0228924
PR	01-SEP-2000	2000US-0229287
PR	01-SEP-2000	2000US-0229343
PR	01-SEP-2000	2000US-0229344
PR	01-SEP-2000	2000US-0229345
PR	03-SEP-2000	2000US-0229509
PR	05-SEP-2000	2000US-0229513
PR	06-SEP-2000	2000US-0230437
PR	06-SEP-2000	2000US-0230438
PR	08-SEP-2000	2000US-0231242
PR	12-SEP-2000	2000US-0231368
PR	14-SEP-2000	2000US-0232397
PR	14-SEP-2000	2000US-0232398
PR	14-SEP-2000	2000US-0232399
PR	14-SEP-2000	2000US-0232400
PR	14-SEP-2000	2000US-0233261
PR	14-SEP-2000	2000US-0233063
PR	14-SEP-2000	2000US-0233064
PR	21-SEP-2000	2000US-0233065
PR	21-SEP-2000	2000US-0233423
PR	21-SEP-2000	2000US-0234774
PR	25-SEP-2000	2000US-0234897
PR	25-SEP-2000	2000US-0234898
PR	26-SEP-2000	2000US-0235844
PR	26-SEP-2000	2000US-0235834
PR	27-SEP-2000	2000US-0235627
PR	29-SEP-2000	2000US-0236527
PR	29-SEP-2000	2000US-0236567
PR	29-SEP-2000	2000US-0236568
PR	29-SEP-2000	2000US-0236569
PR	29-SEP-2000	2000US-0236670
PR	02-OCT-2000	2000US-0236682
PR	02-OCT-2000	2000US-0237038
PR	02-OCT-2000	2000US-0237037
PR	02-OCT-2000	2000US-0237039
PR	02-OCT-2000	2000US-0237040
PR	13-OCT-2000	2000US-0239335
PR	13-OCT-2000	2000US-0239337
PR	20-OCT-2000	2000US-0240960
PR	20-OCT-2000	2000US-0241785
PR	20-OCT-2000	2000US-0241786
PR	20-OCT-2000	2000US-0241787
PR	20-OCT-2000	2000US-0241808
PR	20-OCT-2000	2000US-0244652
PR	08-NOV-2000	2000US-0246475
PR	08-NOV-2000	2000US-0246476
PR	08-NOV-2000	2000US-0246477
PR	08-NOV-2000	2000US-0246478
PR	08-NOV-2000	2000US-0246523
PR	08-NOV-2000	2000US-0246524
PR	08-NOV-2000	2000US-0246525
PR	08-NOV-2000	2000US-0246526
PR	08-NOV-2000	2000US-0246527
PR	08-NOV-2000	2000US-0246528
PR	08-NOV-2000	2000US-0246532
PR	08-NOV-2000	2000US-0246609
PR	08-NOV-2000	2000US-0246610
PR	08-NOV-2000	2000US-0246611
PR	08-NOV-2000	2000US-0246613
PR	17-NOV-2000	2000US-0249207
PR	17-NOV-2000	2000US-0249208
PR	17-NOV-2000	2000US-0249209
PR	17-NOV-2000	2000US-0249210









08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 06-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;  
WPI: 2001-483426/52.

Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
useful for preventing, diagnosing and/or treating cancers and  
metastasis -

Disclosure; SEQ ID NO 25542; 3071bp + Sequence Listing; English.

AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)  
amino acid sequences given in AAM82170 to AAM91821. (I) have cytostatic  
activity, and can be used in gene therapy and vaccine production. (I)  
proteins and polynucleotides may be used in the prevention, diagnosis and  
treatment of diseases associated with inappropriate (I) expression. For  
example, they may be used to treat disorders associated with decreased  
expression by rectifying mutations or deletions in a patient's genome  
that affect the activity of (I) by expressing inactive proteins or to  
supplement the patients own production of (I). Additionally, (I)  
polynucleotides may be used to produce the secreted (I), by inserting  
the nucleic acids into a host cell and culturing the cell to express the  
protein. (I) proteins and polynucleotides may be used to prevent,  
diagnose and treat immune/hematopoietic-related diseases, especially  
cancers and cancer metastases of hematopoietic-derived cells. AAK64703  
to AAK7694 represent human immune/hematopoietic antigen genomic  
sequences from the present invention. AAK54942 to AAK54950 and AAM82169  
represent sequences used in the exemplification of the present invention..

Sequence 2914 BP; 811 A; 676 C; 724 G; 703 T; 0 other;

Query Match 7.1%; Score 39.6; DB 22; Length 2914;  
Best Local Similarity 61.8%; Pred. No. 1.6;  
Matches 63; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

Oy 210 atattatatagtgatatgtaattttccatatatatattacatcaggta 269  
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 409 atatatactaaataatataggtatatacatatatataatttttcataatgta 468

Oy 270 tcccagtcactgtaccatttcccaggagacaatggtgcct 311  
| | | | | | | | | | | | | | | |  
Db 469 ggttaattcagtgtaaatctcccttgggaagaagtgaagct 510

RESULT 14  
ABL34155  
ID ABL34155 standard; DNA: 15548 BP.  
XX  
AC ABL34155;  
XX  
XX 26-MAR-2002 (first entry)  
Dt  
Dx Human immune system associated gene SEQ ID NO: 2128.  
DE  
Dx Human; immune system disease; cytosine methylation; antiasthmatic;  
Kw antihypertensive; antianemic; cytostatic; noctropic;  
Kw neuroprotective; anti-HIV; anticonvulsant; ophthalmological;  
Kw antirheumatic; antiarthritic; antididiabetic; antipsoriatic;  
Kw antinflammatory; cancer; eye disease; arteriosclerosis; anaemia;  
Kw acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
Kw neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;  
Kw gene; ds.  
Xx  
Xx Homo sapiens.  
Os  
Xs WO200200928-A2.  
Pn  
XX  
XX 03-JAN-2002.  
Pd  
XX  
XX 02-JUL-2001; 2001WO-EP07537.  
Pf  
XX  
XX 30-JUN-2000; 2000DE-1032529.  
Pr 01-SEP-2000; 2000DE-1043826.  
XX  
XX  
Pa (EPIC-) EPIGENOMICS AG.  
XX  
Pi Olek A, Piepenbrock C, Berlin K;  
PI WPI; 2002-130909/17.  
DR  
XX  
XX Nucleic acid comprising fragment of chemically modified gene, useful  
Pt for diagnosis and treatment of diseases associated with abnormal  
PT cytosine methylation -  
PS  
XX  
XX Claim 1; SEQ ID NO 2128; 3zpp + Sequence Listing; German.

The present invention provides a number of human immune system associated genes which are modified by the methylation of cytosines. The sequences can be used in the diagnosis and treatment of immune system disorders, including eye diseases such as retinopathy, neovascular glaucoma and macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis, rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel diseases. The present sequence is a gene of the invention.

CC  
Cc  
CC Sequence 15548 BP; 4209 A; 247 C; 2903 G; 8189 T; 0 other;  
Ct  
XX

Query Match 7.0%; Score 39; DB 24; Length 15548;  
Best Local Similarity 71.8%; Pred. No. 5.1;  
Matches 51; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Oy 200 tttagcttcacatattatatagtgatgatgtaattttccatatatatattacatattta 259

[illegible]

RESULT	15
AAQ03743	
ID	AAQ03743 standard; DNA; 2888 BP

Db 1911 ca 1912

Search completed: June 9, 2002, 19:43:45  
Job time: 15669 sec

Query Match	7.0%	Score 38.8	DB 11	Length 2888
Best Local Similarity	47.5%	Pred. No. 2.7		
Matches 115	Conservative 0	Mismatches 127	Indels 0	Gaps 0

[illegible]



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 9, 2002, 12:38:01 ; Search time 7086.26 Seconds  
(without alignments)  
1638.979 Million cell updates/sec

Title: US-09-853-544-1  
Perfect score: 555  
Sequence: 1 ttttttttcttaataac.....agttgtgcagcgtgcacaca 555

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues  
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08  
Maximum Match 1008  
Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_hlg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: gb\_da:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_hlg\_hum:\*  
31: em\_hlg\_inv:\*  
32: em\_hlg\_other:\*  
33: em\_hlgo\_inv:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
------------	-------	-------	-------	--------	----	-------------

C	1	531.6	95.8	4169	6	AX250159	Sequence
C	2	531.6	95.8	14437	2	AC068353	Sequence
C	3	531.6	95.8	158104	2	AF284563	Homo sapi
C	4	518.6	93.4	604	9	HUWYO73E04	Homo sapi
C	5	514.2	92.6	63449	2	AC026976	Homo sapi
C	6	337.4	60.8	174281	2	AC068338	Homo sapi
C	7	265.2	47.8	415	11	G29628	human STS
C	8	101	101	11	G32791	G32791	A009042 Hum
C	9	65.8	11.9	80892	2	AC094509	Rattus no
C	10	63	11.4	2543	9	AK024793	Homo sapi
C	11	63	11.4	5241	6	AX250160	Sequence
C	12	63	11.4	131197	2	AC107883	Homo sapi
C	13	63	11.4	161492	2	AC016693	Homo sapi
C	14	63	11.4	164871	2	AC087465	Homo sapi
C	15	45	8.1	189335	2	AC087477	Homo sapi
C	16	43.8	7.9	190065	2	AC094728	Rattus no
C	17	43.8	7.9	84686	9	AL138685	Human DNA
C	18	43.6	7.9	183037	2	AL139036	Human DNA
C	19	43.4	7.8	189281	2	AL138763	Human DNA
C	20	43.4	7.8	2763	9	AK022387	Homo sapi
C	21	43.2	7.8	32969	9	AY030284	Homo sapi
C	22	43.2	7.8	170071	2	AC084733	Homo sapi
C	23	43.2	7.8	183175	2	AC103564	Homo sapi
C	24	43	7.7	82422	9	AL138763	Human DNA
C	25	42.6	7.7	37393	9	HS023262	Human DNA
C	26	42.6	7.7	142515	9	AC008556	Homo sapi
C	27	42.6	7.7	160656	9	AC024258	Homo sapi
C	28	42.6	7.7	172039	9	AL356741	Human DNA
C	29	42.6	7.7	172264	2	AC099503	Homo sapi
C	30	42.6	7.7	175265	2	AC063964	Homo sapi
C	31	42.6	7.7	185409	2	AC018758	Homo sapi
C	32	42.6	7.7	229162	2	AC097441	Rattus no
C	33	42.4	7.6	155362	2	AC022244	Homo sapi
C	34	42.4	7.6	194181	2	AC018886	Homo sapi
C	35	42.4	7.6	256172	2	AC005139	Plasmodiu
C	36	42.4	7.6	310779	2	AC005140	Plasmodiu
C	37	42.4	7.6	349980	6	AX344569	Sequence
C	38	42.2	7.6	89881	9	AL160264	Human DNA
C	39	42.2	7.6	111547	2	AP002332	Homo sapi
C	40	42.2	7.6	146072	9	AC093828	Homo sapi
C	41	42.2	7.6	152409	2	PFMAL1P1	Plasmodiu
C	42	42.2	7.6	173515	2	AC073394	Homo sapi
C	43	42	7.6	40878	2	AC099066	Homo sapi
C	44	42	7.6	152184	2	AC090630	Homo sapi
C	45	42	7.6	163138	2	AC094481	Rattus no

#### ALIGNMENTS

RESULT	1	AX250159	4169 bp	DNA	linear	PAT 28-SEP-2001
LOCUS	AX250159					
DEFINITION	Sequence	4 from Patent WO0166594.				
ACCESSION	AX250159					
VERSION	AX250159.1	GI:15864534				
KEYWORDS						
SOURCE						
ORGANISM	human.					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
REFERENCE	1 (bases 1 to 4169)					
AUTHORS	Plozman,G.D., Whyte,D., Manning,G., Sudarsanam,S. and Martinez,R.					
TITLE	Human protein kinases and protein kinase-like enzymes					
JOURNAL	Patent: WO 0166594-A 4 13-SEP-2001;					
	Sugen, Inc. (US)					
FEATURES						
Source	1..4169					
	/organism="Homo sapiens"					
	/db_xref="taxon:9606"					
BASE COUNT	871 a	1398 c	1167 g	733 t		
ORIGIN						

```

Query Match          95.8%; Score 531.6; DB 6; Length 4169;
Best Local Similarity 98.9%; Pred. No. 1.3e-130;
Matches 545; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1 ttttttttttaataaacaatttgacaaaagggtgaagaatcctaacaagtattg 60
    |||
DB 4110 tttttttttttttaaataaataatttgacaaaagggtgaagaatcctaacaagtattg 4051

QY 61 aggcacgtgtccacagctgcattcgtcacagaactgtccacagactgttcagactg 120
    |||
DB 4050 AGGCACGTGTCCACAGCTGCATTGCTCACAGAACTCTCTCCAGACGTTGCATGCACT 3991

QY 121 ggaatgtgtataattacagaagaacacagggagagacttagtgcagagagagagagatg 180
    |||
DB 3990 GGAATGTGTATATATTACAGAGAAACAGGAGAGACTTAGTGCAGAGAGAGAGAGAGTG 3921

QY 181 tgaagggagacacagcatccttagtcttcatattatataatgtatattgtatttctata 240
    |||
DB 3930 TGGACGGGCAACGACATCTTAGCTTTCATATTATATGATATGATATTTCATATA 3871

QY 241 tataattatataattatcatccaggtatccacagctatctgtacattccacagagaga 300
    |||
DB 3870 TATATATTATATATTATTAATCATCAGATATCCAGTATCTGTACATTTCCACAGGAGA 3811

QY 301 catgggtgtcttccaaggcagagacagaaaggttaaggcagggagagagcagcaggtgc 360
    |||
DB 3810 CATGGGTGTCTTCCAAAGCGAGACAGAAAGGTTAGCAGAGAGGCGCAGCGATGC 3751

QY 361 aggcgtgggtgtgtgtacagaagctgtcagaagctgtcagaagctgtcagaagctgtcaga 420
    |||
DB 3750 AGGCTGGGCTTGGCTACAGAAAGCTGCAGAGAGCTTACGACGTGTAGAGAGGCGCCGGG 3691

QY 421 ctccgcagacgcagagctact-gagncacaaagccagctcctcagctccacgcccgcctgcg 479
    |||
DB 3690 CTCGCGACGACGCCAGTACTGGGACGACAAACCAAGCTCTCCAGTCCACGCCGCCGCTGCG 3631

QY 480 atccacgccttctccgcaaatcatcatcatcagagggccgccttcatgtcatcagtt 539
    |||
DB 3630 ATCCACGCCCTTCTCCGCAAACTTCATCATCATCAGGCGCCGCTTCATGTCAGTTCAGTT 3571

QY 540 gtccagcgtgc 550
    |||
DB 3570 GTCCAGCGTGC 3560

RESULT 2
AC068353 144327 bp DNA linear HTG 13-FEB-2002
LOCUS Homo sapiens chromosome 8 clone RP11-399J23 map 8, ** SEQUENCING
DEFINITION IN PROGRESS ***, 3 unordered pieces.
ACCESSION AC068353
VERSION AC068353.11 GI:18653561
KEYWORDS HTG: HTGS_PHASE1; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 144327)
REFERENCE 1 (bases 1 to 144327)
TITLE Homo sapiens chromosome 8, clone RP11-399J23
JOURNAL Unpublished
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguski,K., Boucknight,B., Brown,A., Burkett,G.,
Campione,A., Castle,A., Choedel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dearlano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domingo,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Glade,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hages,B., Heathford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,T., Lacombe,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., MacDonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheters,R.,
Meldrum,J., Meneau,L., Mihova,T., Miranda,C., Mienga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,K., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Testafaye,S., Theodore,J., Tirrell,A., Travers,M., Trigglio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (02-MAY-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Feb 13, 2002 this sequence version replaced g1:18583956.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L10145
Center clone name: 399_J_23
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 109580: contig of 109580 bp in length
* 109581 109680: gap of 100 bp
* 109681 142410: contig of 32730 bp in length
* 142411 142510: gap of 100 bp
* 142511 144327: contig of 1817 bp in length.
Location/Qualifiers
1.144327 "Homo sapiens"
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8"
/clone="RP11-399J23"
/clone_lib="RPCT-11 Human Male BAC"
BASE COUNT 41798 a 30791 c 30089 g 41448 t 201 others
ORIGIN
Query Match          95.8%; Score 531.6; DB 2; Length 144327;
Best Local Similarity 98.9%; Pred. No. 2.9e-130;
Matches 545; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1 ttttttttttaataaacaatttgacaaaagggtgaagaatcctaacaagtattg 60
    |||
DB 5608 TTTTTTTCCTTAATTAATTAACAAATTTGACAAAGGTCGAACAATCTTAACAAGTATTG 56067

QY 61 aggcacgtgtccacagctgcattcgtcacagaactgtccacagactgttcagactg 120
    |||
DB 5608 AGGCACGTGTCCACAGCTGCATTGCTCACAGAACTCTCTCCAGACGTTGCATGCACT 56127

QY 121 ggaatgtgtataattacagaagaacacagggagagacttagtgcagagagagagagatg 180
    |||
DB 56128 GGAATGTGTATATATTACAGAGAAACAGGAGAGACTTAGTCACAGAGAGAGAGAGTG 56187

QY 181 tgaagggagacacagcatccttagtcttcatattatataatgtatattgtatttctata 240
    |||
DB 56188 TGGACGGGCAACGACATCTTAGCTTTCATATTATGATATGATATTTCATATA 56247

QY 241 tataattatataattatcatccaggtatccacagctatctgtacattccacagagaga 300

```

```
|||||
Db 56248 TATATATTATATATTATTTACACGAGTATCCAGTATCTGTACCATTTCCAGGAGCA 56307
OY 301 catggtgtcttccaaagcgagacaagaaaggttaggcaggaagggagcagcagctgc 360
Db 56308 CATGGGTGCTTCCAAAGGCGACAGCAAGAAAGGTTAGCGAGGAGGAGGCGACGAGCTGC 56367
OY 361 aggtctggggtctgtgtctcacaagagctgcagagcttcagcagctglaaagaagcccgag 420
Db 56368 AGGCTGGGGCTTGCTCACAGAAAGCTGCAGAGCTTCACGCACTGTAAAGAGGGCCCGGG 56427
OY 421 ctccgcaagcagcagcagctact-gagmcaagcagctctcctcagcagcccgctgcg 479
Db 56428 CTCGCCAGACGCGAGGTAAGTGGACAAAGCCAGTCTCAGCTCCAGCCCGCTGCG 56487
OY 480 atccacgcctctcgcagaaactcatcatcatcagagcgccgctcatgcatcagct 539
Db 56488 ATCCACGCCCTTCTCGCAAAATTCATCATCATCAGAGCCCGCTTCATCATCAGT 56547
OY 540 gtccagcagctgc 550
Db 56548 GTCCAGCGCTGC 56558
```

```
RESULT 3
AF284563/c 158104 bp DNA linear HTG 13-JUN-2001
LOCUS Homo sapiens chromosome 8 clone CTC-776F21 map 8p22-p23, WORKING
DEFINITION DRAFT SEQUENCE, 8 unordered pieces.
ACCESSION AF284563
VERSION AF284563.2 GI:14389395
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 158104)
Reichwald, K., Baumgart, C., Blechschmidt, K., Dette, M., Jahn, N.,
Lehmann, R., Menzel, U., Polley, A., Schlöbel, M. B., Schudy, A.,
Siddiqui, R., Taudien, S., Wen, G., Rosenthal, A. and Platzer, M.
Chromosome 8 genomic sequence
Unpublished
2 (bases 1 to 158104)
Genome Sequencing Center Jena.
Direct Submission
Submitted (04-JUN-2000) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
On Jun 13, 2001 this sequence version replaced gi:9082332.
----- Genome Center
Center: Institute of Molecular Biotechnology
Center code: IMB
Web site: http://genome.imb-jena.de/
Contact: gscj-submit@genome.imb-jena.de
----- Project Information
Center project name: H475
Center clone name: CTC-776F21
----- Summary Statistics
Sequencing vector: M13; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 147627 bases at least Q40
Consensus quality: 151667 bases at least Q30
Consensus quality: 154312 bases at least Q20
Quality coverage: 4.50 x in Q20 bases; sum-of-contigs
```

Sequence Quality Assessment:  
This entry has been annotated with sequence quality  
estimates computed by the Phrap assembly program.  
All manually edited bases have been reduced to quality 10.  
Quality levels above 40 are expected to have less than  
1 error in 10,000 bp.  
Base-by-base quality values are not generally visible from the  
GenBank flat file format but are available as part

of this entry's ASN.1 file.

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 8 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

```
1 2777: contig of 2777 bp in length
* 2778 2877: gap of unknown length
* 2878 42406: contig of 39529 bp in length
* 42407 42506: gap of unknown length
* 42507 78441: contig of 35935 bp in length
* 78442 78541: gap of unknown length
* 78542 82141: contig of 3599 bp in length
* 82141 82240: gap of unknown length
* 82241 115504: contig of 33264 bp in length
* 115505 115604: gap of unknown length
* 115605 137507: contig of 21903 bp in length
* 137508 137607: gap of unknown length
* 137608 150985: contig of 13378 bp in length
* 150986 151085: gap of unknown length
* 151086 158104: contig of 7019 bp in length.
```

## FEATURES

## source

1..158104

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="8"

/map="8p22-p23"

/clone="CTC-776F21"

## misc\_feature

1..4

/note="assembly-fragment"

clone\_end:SP6

vector\_side:left

158101..158104

/note="assembly-fragment"

clone\_end:T7

vector\_side:right

## BASE COUNT

44317 a 35347 c 34418 g 43322 t 700 others

## ORIGIN

## Query Match

Best Local Similarity 95.8%; Score 531.6; DB 2; Length 158104;

Matches 545; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

```
OY 1 ttttttcttaataaatttgcacaaagggaggaagaaatcctaagaagratatg 60
Db 66956 TTTTCTTCTTTAAATTAATTAACAAATTTGACAAAGGTAAGAAATCTTAAACAAGGATTTG 66897
OY 61 aggcagtgctccagcagctgcatcagttcacagaacgctccagagcgttcagtaact 120
Db 66896 AGGCAGTGTCCAGCGCTGCATTCAGTTCACAGAACTGTCTCAGACGTTGCACTGGAAC 66837
OY 121 ggaatgtgtataataatcacagaagaaacacggaggaactagtcagagagagagcagtc 180
Db 66836 GGAATGTGTATTAATTAACAGAAAGGAGGAGGAGCTTACAGAGAGAGACAGAGTG 66777
OY 181 tggacgggaaacagcagcctctagctcttcatattatataatgatatatatttata 240
Db 66776 TGGACGGGAAACGCAATCTTACGCTTCTTCAATTTATATATATGATATGATTTCTATA 66717
OY 241 tatatatatatatttcatcaggtatccagcagcatctgtacatctccagggaga 300
Db 66716 TATATTTTATATATTTTATTCATCCAGGTATCCAGTATCTGTATTCATTTCCAGGAGA 66657
OY 301 catggtgtcttccaaagcgagacaagaaaggttaggcaggaagggagcagcagcgtgc 360
Db 66656 CATGGGTGCTTCCAAAGGCGACAGCAAGAAAGGTTAGGAGGAGGAGGCGACGAGCTGC 66597
OY 361 aggtctggggtctgtgtctcacaagagcgtgcaggaagcttcagcagcagctgtaagaggcccg 420
```

|||||  
Db 66596 AGGCTGGGCTTGCTACAGAGCTCAGAGAGCTGTCAAGAGGCGCCCGG 66537  
OY 421 CTCGCGAGCGCAGTCT-9agunaaagccagtcctcagccacgcccgcgcg 479  
Db 66536 CTCGCGAGAGCCAGTACTGCGAGCAAGCCACTCTCCAGCTCCAGCCCGCTGCG 66477  
OY 480 atccaccgctctccgaaactcattcattcattcagggccgcgtctcagtcagtcagtc 539  
Db 66476 ATCCAGCGCTTCTCCGCAAACTTCATCATCATCAGGCGCCGCTTCATGTGATCCAGTT 66417  
OY 540 gtgcagcgctgc 550  
Db 66416 GTGCAGCGTGC 66406

RESULT 4  
HUM073E04/c 604 bp mRNA linear PRI 04-AUG-1998  
LOCUS HUM073E04  
DEFINITION Homo sapiens full length insert cDNA Y073E04.  
ACCESSION AF075060  
VERSION AF075060.1 GI:3377601  
KEYWORDS FLI\_CDNA.  
SOURCE human.  
ORGANISM Homo sapiens  
Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE 1 (bases 1 to 604)  
AUTHORS Moessner,J., Tan,F., Marra,M., Kucaba,T., Yandell,M., Martin,J.,  
Marth,G., Bowles,L., Wylie,T., Bowers,Y., Steptoe,M., Theisling,B.,  
Geisel,S., Allen,M., Underwood,K., Chappell,J., Person,B.,  
Gibbons,M., Harvey,N., Page,D., Chamberlain,A., Morales,R.,  
Schuck,R., Ritzer,E., Kohn,S., Swaller,T., Behner,K., Hillier,L.,  
Wilson,R. and Waterston,R.  
TITLE Full Clone Sequencing of the Longest Available Member from Each  
Unigene Cluster  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 604)  
AUTHORS Waterston,R.  
TITLE Direct Submission  
JOURNAL Submitted (30-JUN-1998) Department of Genetics, Washington  
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
COMMENT SUBMITTED BY: Genome Sequencing Center  
Department of Genetics  
Washington University  
St. Louis MO 63108, USA  
http://genome.wustl.edu/gsc  
mailto:estewatson.wustl.edu

NOTICE: This sequence represents the full insert of this cDNA. No  
attempt has been made to verify whether this corresponds to the  
full-length of the original mRNA from which it was derived. We  
have tried to obtain double-stranded, or double chemistry sequence  
across the entire clone, but potentially, there are areas in the  
sequence where this level of coverage was not achieved.  
Nevertheless, we are confident of the accuracy of this sequence as  
all regions of low quality, as defined by PHRAP (P. Green, in  
preparation), were visually inspected and edited accordingly. The  
consensus quality values for this sequence have been submitted  
separately.

FEATURES  
source The location of this clone is unknown.  
1.604  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:183582"  
/clone\_lib="Soares\_breast\_3MNBHst"

BASE COUNT 128 a 164 c 159 g 153 t

Query Match 93.4%; Score 518.6; DB 9; Length 604;  
Best Local Similarity 98.9%; Pred. No. 2,5e-127;  
Matches 532; Conservative 0; Mismatches 5; Indels 1; Gaps 1;  
OY 14 aaataaattttgacaaaaggtggaagaattcttaacacagatgtgagccagttcca 73  
Db 604 AAATTAACAAATTTGACAAAAGGGTGAAGAAATCTTAAACAAGGATATGAGCCAGTCCCA 545  
OY 74 ggcgtcattcagttcaacagacgtctcctcagagcgctgtgcatggaactggaatgtata 133  
Db 544 GGCTGCAATTCAGTTCCAGACAGACTGTCTCTCAGAGCGTTTGCACTGGAAATGTGTATA 445  
OY 134 attacagaagaacacagaggaagcattagtcagaagaagagacagtgctgagcggacaca 193  
Db 484 ATTACAGAAACAAAACAGGAGGAGCTTAGTCAGAGAGACAGACAGAGTGTGACGGCAACA 425  
OY 194 gcatccttagctttcattatattatattatattatattatattatattatattatattat 253  
Db 424 GCATCTTACTCTTTTAT 365  
OY 254 atttaacatccagttatccagatcctcagtcattcctccagggagacatggtgtcttc 313  
Db 364 ATTTTACATCCAGGTATCCAGTCAATCTGTACCATTTCCAGGAGACATGGGTCTTCC 305  
OY 314 aagcgagacaggaagaggtgtaagcagaggaagggcagcagtgcaaggtctgggcttg 373  
Db 304 AAGCGAGACAGAAAGGGTTAGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 245  
OY 374 gcttcacagaagcgtgcagaggtcttcagcagctgtaagaagggcccggtctcgaagagcc 433  
Db 244 GCTCACGAAGGCTGACGAGACTTTCAGCGATGTAAAGGGGCCCGGCTCCGACAGGCC 185  
OY 434 aggtact-9agunaaagccagtcctcagtcacagcccgccgtcagtcacagccgttc 492  
Db 184 AGGTACTGCGAGCAAAAGCCAGTCTCAGCTCCAGCTCCAGCCCGCCGTGATACAGCGCTTC 125  
OY 493 tccgcaacttcattcattcattcattcattcattcattcattcattcattcattcattcatt 550  
Db 124 TCCGCAAACTTCATCATCATCAGGCGCCGCTTCATGTGATTCAGTGTGTCAGCGTGC 67

RESULT 5  
LOCUS AC026976 63449 bp DNA linear HTG 24-MAR-2001  
DEFINITION Homo sapiens clone RP11-792016, LOW-PASS SEQUENCE SAMPLING.  
ACCESSION AC026976  
VERSION AC026976.2 GI:13443237  
KEYWORDS HTG; HTGS\_PHASE0.  
SOURCE human.  
ORGANISM Homo sapiens  
Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE 1 (bases 1 to 63449)  
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
TITLE Unpublished  
JOURNAL Homo sapiens, clone RP11-792016  
REFERENCE 2 (bases 1 to 63449)  
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Bede,F.,  
Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G.,  
Campiano,A., Castle,A., Choquet,Y., Colangelo,M., Collins,S.,  
Collamore,A., Cooke,P., DeRubeis,K., Dewar,K., Diaz,J.S.,  
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,N., Gage,D.,  
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,  
Grand-Pierre,N., Grant,G., Hagos,B., Heathford,A., Horton,L.,  
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kamp,L., Karatas,A.,  
Klein,U., Lacombe,K., Lamazares,R., Lander,T., Lehotzky,J.,  
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,  
McCarthy,M., McEwan,P., McGuirk,A., McKernan,K., McPheters,R.,  
Melgrim,J., Meneses,L., Mihova,T., Mironov,C., Mienga,Y., Morrow,J.,  
Murphy,T., Naylor,T., Norman,C.H., O'Connor,T., O'Donnell,P.,  
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,  
Pisanil,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,

TITLE  
JOURNAL

## COMMENT

Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,  
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,  
Testaye, S., Theodore, J., Tirrell, A., Travers, M., Triggillo, J.,  
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,  
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Submitted (25-MAR-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Mar 24, 2001 this sequence version replaced g1:7328855.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project Information

Center project name: 18651

Center clone name: 792\_O\_16

-----  
\* NOTE: This record contains 78 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.  
\* 1  
\* 707 806: gap of 100 bp in length  
\* 807 1521: contig of 715 bp in length  
\* 1522 1621: gap of 100 bp  
\* 1622 2316: contig of 695 bp in length  
\* 2317 2416: gap of 100 bp  
\* 2417 3136: contig of 720 bp in length  
\* 3137 3236: gap of 100 bp  
\* 3237 3939: contig of 703 bp in length  
\* 3940 4039: gap of 100 bp  
\* 4040 4731: contig of 692 bp in length  
\* 4732 4831: gap of 100 bp  
\* 4832 5557: contig of 726 bp in length  
\* 5558 5657: gap of 100 bp  
\* 5658 6378: contig of 721 bp in length  
\* 6379 6478: gap of 100 bp  
\* 6479 7201: contig of 723 bp in length  
\* 7202 7301: gap of 100 bp  
\* 7302 8007: contig of 706 bp in length  
\* 8008 8107: gap of 100 bp  
\* 8108 8890: contig of 783 bp in length  
\* 8891 8990: gap of 100 bp  
\* 8991 9661: contig of 671 bp in length  
\* 9662 9761: gap of 100 bp  
\* 9762 10426: contig of 665 bp in length  
\* 10427 10526: gap of 100 bp  
\* 10527 11248: contig of 722 bp in length  
\* 11249 11348: gap of 100 bp  
\* 11349 12061: contig of 713 bp in length  
\* 12062 12161: gap of 100 bp  
\* 12162 12857: contig of 696 bp in length  
\* 12858 12957: gap of 100 bp  
\* 12958 13669: contig of 712 bp in length  
\* 13670 13769: gap of 100 bp  
\* 13770 14542: contig of 773 bp in length  
\* 14543 14642: gap of 100 bp  
\* 14643 15348: contig of 706 bp in length  
\* 15349 15448: gap of 100 bp  
\* 15449 16159: contig of 711 bp in length  
\* 16160 16259: gap of 100 bp  
\* 16260 16998: contig of 739 bp in length

15999 17098: gap of 100 bp  
17099 17810: contig of 712 bp in length  
17811 17910: gap of 100 bp  
17911 18631: contig of 721 bp in length  
18632 18731: gap of 100 bp  
18732 19430: contig of 699 bp in length  
19431 19530: gap of 100 bp  
19531 20246: contig of 716 bp in length  
20247 20346: gap of 100 bp  
20347 21068: contig of 722 bp in length  
21069 21168: contig of 100 bp  
21169 21895: contig of 727 bp in length  
21896 21995: gap of 100 bp  
21996 22714: contig of 719 bp in length  
22715 22814: gap of 100 bp  
22815 23536: contig of 722 bp in length  
23537 23636: gap of 100 bp  
23637 24355: contig of 719 bp in length  
24356 24455: gap of 100 bp  
24456 25152: contig of 697 bp in length  
25153 25252: gap of 100 bp  
25253 25957: contig of 705 bp in length  
25958 26057: gap of 100 bp  
26058 26743: contig of 686 bp in length  
26744 26843: gap of 100 bp  
26844 27559: contig of 716 bp in length  
27560 27659: gap of 100 bp  
27660 28384: contig of 725 bp in length  
28385 28484: gap of 100 bp  
28485 29195: contig of 711 bp in length  
29196 29295: gap of 100 bp  
29296 30012: contig of 717 bp in length  
30013 30112: gap of 100 bp  
30113 30811: contig of 699 bp in length  
30812 30911: gap of 100 bp  
30912 31612: contig of 701 bp in length  
31613 31712: gap of 100 bp  
31713 32426: contig of 714 bp in length  
32427 32526: gap of 100 bp  
32527 33233: contig of 707 bp in length  
33234 33333: gap of 100 bp  
33334 34090: contig of 757 bp in length  
34091 34190: gap of 100 bp  
34191 34937: contig of 747 bp in length  
34938 35037: gap of 100 bp  
35038 35748: contig of 711 bp in length  
35749 35848: gap of 100 bp  
35849 36564: contig of 716 bp in length  
36565 36664: gap of 100 bp  
36665 37370: contig of 706 bp in length  
37371 37470: gap of 100 bp  
37471 38166: contig of 696 bp in length  
38167 38266: gap of 100 bp  
38267 38985: contig of 719 bp in length  
38986 39085: gap of 100 bp  
39086 39808: contig of 723 bp in length  
39809 39908: gap of 100 bp  
39909 40632: contig of 724 bp in length  
40633 40732: gap of 100 bp  
40733 41460: contig of 728 bp in length  
41461 41560: gap of 100 bp  
41561 42307: contig of 747 bp in length  
42308 42407: gap of 100 bp  
42408 43119: contig of 712 bp in length  
43120 43219: gap of 100 bp  
43220 43954: contig of 735 bp in length  
43955 44054: gap of 100 bp  
44055 44770: contig of 716 bp in length  
44771 44870: gap of 100 bp  
44871 45593: contig of 723 bp in length  
45594 45693: gap of 100 bp  
45694 46403: contig of 710 bp in length  
46404 46503: gap of 100 bp

```
* 46504 47213: contig of 710 bp in length
* 47214 47313: gap of 100 bp
* 47314 48028: contig of 715 bp in length
* 48029 48128: gap of 100 bp
* 48129 48825: contig of 697 bp in length
* 48826 48925: gap of 100 bp
* 48926 49596: contig of 671 bp in length
* 49597 49696: gap of 100 bp
* 49697 50397: contig of 701 bp in length
* 50398 50497: gap of 100 bp
* 50498 51211: contig of 714 bp in length
* 51212 51311: gap of 100 bp
* 51312 52032: contig of 721 bp in length
* 52033 52132: gap of 100 bp
* 52133 52851: contig of 719 bp in length
* 52852 52951: gap of 100 bp
* 52952 53672: contig of 721 bp in length
* 53673 53772: gap of 100 bp
* 53773 54467: contig of 695 bp in length
* 54468 54567: gap of 100 bp
* 54568 55275: contig of 708 bp in length
* 55276 55375: gap of 100 bp
* 55376 56072: contig of 697 bp in length
* 56073 56172: gap of 100 bp
```

Query Match 92.6%; Score 514.2; DB 2; Length 63449;  
Best Local Similarity 98.0%; Pred. No. 1e-125;  
Matches 541; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

```
QY 1 ttttttttcttaataacaaatttgcacaaagggtggaagaatccttaacaaagctatg 60
    |||
DB 49030 ttttttttcttaataacaaatttgcacaaagggtggaagaatccttaacaaagctatg 49089
    |||
QY 61 aggcagctgtccagctgctcagctcagctcagctcagctcagctcagctcagctcagct 120
    |||
DB 49090 aggcagctgtccagctgctcagctcagctcagctcagctcagctcagctcagctcagct 49149
    |||
QY 121 ggaatgtgtataataacaaagggtggaagaatccttaacaaagctatg 180
    |||
DB 49150 ggaatgtgtataataacaaagggtggaagaatccttaacaaagctatg 49209
    |||
QY 181 tgaagcggaacagcagctcagctcagctcagctcagctcagctcagctcagctcagct 240
    |||
DB 49210 tgaagcggaacagcagctcagctcagctcagctcagctcagctcagctcagctcagct 49269
    |||
QY 241 tatatatatatattatcatcagctcagctcagctcagctcagctcagctcagctcagct 300
    |||
DB 49270 tatatatatatattatcatcagctcagctcagctcagctcagctcagctcagctcagct 49329
    |||
QY 301 catgggtgtcttccaagcgagacaggaagggttaagcgaggaaggcgagcgagctg 360
    |||
DB 49330 catgggtgtcttccaagcgagacaggaagggttaagcgaggaaggcgagcgagctg 49389
    |||
QY 361 agagctgtgagctgtgctcacaagaagctgctcagctcagctcagctcagctcagctcagct 419
    |||
DB 49390 agagctgtgagctgtgctcacaagaagctgctcagctcagctcagctcagctcagctcagct 49449
    |||
QY 420 gctccgcaagcgagctgact-gagncaaagcagctcctcagctcagctcagctcagctcagct 478
    |||
DB 49450 gctccgcaagcgagctgact-gagncaaagcagctcctcagctcagctcagctcagctcagct 49509
    |||
QY 479 gatccacgagccttctccgaaactcatcatcagctcagctcagctcagctcagctcagctcagct 538
    |||
DB 49510 gatccacgagccttctccgaaactcatcatcagctcagctcagctcagctcagctcagctcagct 49569
    |||
QY 539 tgtgcagcgctgc 550
    |||
DB 49570 tgtgcagcgctgc 49581
    |||
```

RESULT 6  
AC068338 AC068338 174281 bp DNA linear HTG 01-DEC-2001  
LOCUS

DEFINITION Homo sapiens chromosome 11 clone RP11-817013 map 11, WORKING DRAFT  
SEQUENCE, 12 unordered pieces.  
AC068338  
AC068338.5 GI:17223174  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS

1 (bases 1 to 174281)  
Birken, B., Linton, L., Nusbaum, C. and Lander, E.  
Homo sapiens chromosome 11, clone RP11-817013  
Unpublished  
2 (bases 1 to 174281)  
Birken, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bede, F.,  
Boguslavsky, L., Bouknight, B., Brown, A., Burkett, G.,  
Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,  
Collins, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.S.,  
Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,  
Galagan, J., Gardyna, S., Glende, S., Goyette, M., Graham, L.,  
Grand-Pierre, N., Grant, G., Hags, B., Heatford, A., Horton, L.,  
Howard, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karakas, A.,  
Klein, J., Lacombe, K., Lamazares, R., Landers, T., Lehotzky, J.,  
Levine, R., Liu, C., Liu, G., Locke, K., MacDonald, P., Margulis, N.,  
McCarthy, M., McEwan, P., McGuire, A., McKernan, K., McPheters, R.,  
Melrim, J., Menus, L., Mihova, T., Miskern, C., Mlenka, V., Morrow, J.,  
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,  
O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,  
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,  
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,  
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,  
Tessier, S., Theodore, J., Tirrell, A., Travers, M., Triggillo, J.,  
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,  
Young, G., Zahoun, J., Zimmer, A. and Zody, M.

TITLE  
JOURNAL  
COMMENT

Submitted (02-MAY-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Dec 1, 2001 this sequence version replaced gi:15290773.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIRB  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence.submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L10087  
Center clone name: 817\_O\_13  
----- Summary Statistics  
Sequencing vector: M13; M7815; 0% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 167722 bases at least Q40  
Consensus quality: 169015 bases at least Q30  
Consensus quality: 170603 bases at least Q20  
Insert size: 177000; agarose-fp  
Insert size: 173181; sum-of-contigs  
Quality coverage: 12.6 in Q20 bases; agarose-fp  
Quality coverage: 12.9 in Q2.

NOTE: This is a 'working draft' sequence. It currently  
consists of 12 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as  
runs of N, but the exact sizes of the gaps are unknown.  
This record will be updated with the finished sequence  
as soon as it is available and the accession number will  
be preserved.

1 107104: contig of 107104 bp in length  
\* 107105 107204: gap of 100 bp  
\* 107205 107765: contig of 561 bp in length

```

* 10766 107865: gap of 100 bp
* 107866 108310: contig of 445 bp in length
* 108311 108410: gap of 100 bp
* 108411 109129: contig of 719 bp in length
* 109130 109229: gap of 100 bp
* 109230 109871: contig of 642 bp in length
* 109872 109971: gap of 100 bp
* 109972 110628: contig of 657 bp in length
* 110629 110728: gap of 100 bp
* 110729 112287: contig of 1559 bp in length
* 112288 112387: gap of 100 bp
* 112388 113548: contig of 1161 bp in length
* 113549 113648: gap of 100 bp
* 113649 115950: contig of 2302 bp in length
* 115951 116050: gap of 100 bp
* 116051 119604: contig of 3554 bp in length
* 119605 119704: gap of 100 bp
* 119705 135629: contig of 15925 bp in length
* 135630 135728: gap of 100 bp
* 135730 174281: contig of 38552 bp in length.

```

FEATURES  
source  
1. 174281  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="11"  
/map="11"  
/clone="RP11-817013"  
/clone\_lib="RPC1-11 Human Male BAC"  
1. 107104

```

misc_feature /note="assembly-fragment"
clone_end:SP6
vector_side:left"
misc_feature /note="assembly-fragment"
107205. 107765
misc_feature /note="assembly-fragment"
107866. 108310
misc_feature /note="assembly-fragment"
108411. 109129
misc_feature /note="assembly-fragment"
109230. 109871
misc_feature /note="assembly-fragment"
109972. 110628
misc_feature /note="assembly-fragment"
110729. 112287
misc_feature /note="assembly-fragment"
112388. 113548
misc_feature /note="assembly-fragment"
113649. 115950
misc_feature /note="assembly-fragment"
116051. 119604
misc_feature /note="assembly-fragment"
119705. 135629
misc_feature /note="assembly-fragment"
135730. 174281
misc_feature /note="assembly-fragment"
clone_end:T7
vector_side:right"

```

BASE COUNT 42539 a 44506 c 43805 g 42321 t 1110 others  
ORIGIN

Query Match 60.8%; Score 337.4; DB 2; Length 174281;  
Best Local Similarity 90.8%; Pred. No. 1.5e-78;  
Matches 435; Conservative 0; Mismatches 35; Indels 9; Gaps 7;

```

OY 77 tgcattcagtcacagaactgtccctcagagacgttgatgaatgaatgtgtataatt 136
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 107212 TGCATTTCAGTCACAGAACTGTCGCCAGACGTTGCACTGCAACTGGAATGTATTAA-T 107270

OY 137 acagaaagaaacagagagagactagtcagagagagagacagtgctgacgagcaacagca 196
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 107271 ACAGAAAGAAACAGAGAGAGCTTAGTGACAGAGAGAGAGAGAGTGTGACGGCAACAGCA 107330

OY 197 tccctagcttcattcattatataatgtgtatataatgtatatttcctatataattatata-- 254

```

```

||||| ||||||| ||||||| ||| | ||| |
DB 107331 TCCTTAGTCTTCA-ATTTATATGATATGATTTTGAAGTGAAGCTTCANTNTA 107389
OY 255 -ttcacacagagatccacagcatcgttaccac-tttccagagagagatggtgtctc 312
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 107390 GTTTTACATCCAGGATCCAGTCACTTATCTTACCACTTTCCAGAGGAGATGGGCTTTC 107449
OY 313 caagcgagacagagaaaggttagcagagagagagagagagagagagagagagagagagag 372
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 107450 CAAGCGAGACAGAGAAAGGTTATAGCGAGGAGAGGAGGAGAGAGAGAGAGAGAGAGAG 107509
OY 373 ggcctcacaagctcagagagagagagagagagagagagagagagagagagagagagagag 432
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 107510 GGCCTC-CAGAGAGCTGCAGAGAGCTTC-GCAGACTGTAAAGAGGAGGAGGAGGAGGAG 107567
OY 433 caggtact-gagucaaagcagctcctcagctcagctcagctcagctcagctcagctcagctc 491
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 107568 CAGGTACTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 107627
OY 492 ctccgcaacttcatacatcagagagagagagagagagagagagagagagagagagagagag 550
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 107628 CTCGCAAACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 107686

```

RESULT 7  
G29628 415 bp DNA linear STS 05-OCT-1996  
LOCUS human STS SHGC-34242, sequence tagged site.  
DEFINITION G29628  
ACCESSION G29628.1 GI:1593179  
VERSION G29628.1 GI:1593179  
KEYWORDS STS; STS sequence; primer; sequence tagged site.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 415)  
REFERENCE  
AUTHORS Myers,R.M.  
JOURNAL Unpublished (1996)  
COMMENT  
Contact: Richard M. Myers  
Stanford Human Genome Center (SHGC)  
Stanford University School of Medicine  
Department of Genetics, W-344, Stanford, CA 94305, USA  
Tel: 4157259687  
Fax: 4157259689  
Email: myers@shgc.stanford.edu

Primer A: TTGACAAAGGTCGAGAGAA  
Primer B: TCGTCTGCACTAGTCCTCC  
STS size: 150  
PCR Profile:  
Initial incubation: 94 degrees C for 90 seconds  
Denaturation: 94 degrees C for 15 seconds  
Annealing: 62 degrees C for 23 seconds  
Polymerization: 72 degrees C for 30 seconds  
PCR Cycles: 30  
Thermal Cycler: Perkin Elmer 9600

Protocol:  
Template: 25 ng  
Primer: each 1 uM  
dNTPs: each 200 uM  
Tag Polymerase: 0.05 units/uM  
Total Vol: 10 uL

Buffer:  
MgCl2: 2.5 mM  
KCl: 50 mM  
Tris-HCl: 20 mM  
pH: 8.3

Prepared with primer pairs provided by Sandoz, derived from H44044  
-- Washington University/Merck EST sequence.

FEATURES  
source 1. .415 Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="8"  
STS  
primer\_bind 1. .150  
primer\_bind 1. .21  
complement(129. .150)  
BASE COUNT 106 a 84 c 120 g 95 t 10 others  
ORIGIN

Query Match 47.8%; Score 265.2; DB 11; Length 415;  
Best Local Similarity 86.7%; Pred. No. 6.7e-60;  
Matches 357; Conservative 0; Mismatches 43; Indels 12; Gaps 6;

QY 23 ttgacaaaagggtgagaacatcttaacaaggtattgagcgagctgcagctcatt 82  
|||||  
Db 1 TTTGACAAAAGGCTGAGAAATCTTAAACAGGTATTGAGGCCAGTGTCCAGCTGCATT 60  
|||||

QY 83 cagttcacagaactgtctcgaagcgttcgatgaaactggaaatgtataatcagaa 142  
|||||  
Db 61 CACTTCACAGAACTGCTCAGCAGCGTGCATGGAAGTGAATGTATATTNCAGAA 120  
|||||

QY 143 gaaacagaggaagacttagtcagagaagagcagctgtgagcggaacagcatcc 202  
|||||  
Db 121 GAAACAGGAGAGCTTAGTGCAGAGAGGAGAGAGTGTGAGCGGCAACACATCTTA 180  
|||||

QY 203 gtcttcataattatatagtatgatatcttctataatataatataatattacat 262  
|||||  
Db 181 GTCTTCATATTATATATGCTATATGCTATTTCTTATATATATATATATATATAT 240  
|||||

QY 263 cca-ggtatcccaagctcgtacatccattccca-ggagacatgggtgcttccaaagcg 320  
|||||  
Db 241 CCNGGGNNTCCTCAGTCATCTGATTCATNNCCAGGAGAGAGCGGCTCTTCAAGGGC 300  
|||||

QY 321 gaaa-ggtaaggttagcgaggg---aagggcgagcgagcggtgcagctgggctt-- 372  
|||||  
Db 301 GAACACAGGAAAGGTTAGGCCCAAGGACAGGAGGCGAGCGGTCCAGGCTGGGGNATTG 360  
|||||

QY 373 -ggctcacgaagctgcagagagc-ttcagcagctgtcaagagggcccgagct 422  
|||||  
Db 361 GGCTNACAGAACCTTCAGGGGCTTTCAGACCTTTTAAGGGGGCCCCCGCT 412  
|||||

RESULT 8  
G32791 101 bp DNA linear STS 24-SEP-1999  
LOCUS G32791  
DEFINITION A009042 Human Homo sapiens STS genomic, sequence tagged site.  
ACCESSION G32791  
VERSION G32791.1 GI:5923312  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 101)  
AUTHORS Adams M.D.  
TITLE Human STS sequences  
JOURNAL Unpublished  
COMMENT  
Contact: Mark Adams  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850  
Email: mdadams@tigr.org  
Primer A: ATGGAAGTGAAGTGTATATAT  
Primer B: ATATGAAAGACTTAAGATGCTGT  
STS size: 101  
PCR Profile:  
Denaturation: 96C 5min  
Anneal: 54C 30sec  
Extend: 72C 30sec  
Denature: 95C 30sec

FinalExtend: 72C 5min  
Cycles: 30  
Protocol:  
GenomicDNA: 25 ng  
Primer: 0.43 uM each  
dNTPs: 230 uM each  
Ampliflag: 0.5 units  
TagStart Ab: 0.5 units  
Total Volume: 10 ul

Buffer:  
Tris-HCl pH8.8: 100 mM  
KCl: 500 mM  
MgCl2: 20 mM  
Tris-HCl pH8.8: 100 mM  
Concentration: 10X

Prepared with primer pairs derived from THC129699: GenBank  
Accession Numbers: R70141, H44044, L73763.  
Location/Qualifiers  
1. .101  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="Human"  
STS  
primer\_bind 1. .101  
primer\_bind 1. .23  
complement(79. .101)  
BASE COUNT 35 a 14 c 30 g 22 t  
ORIGIN

Query Match 18.2%; Score 101; DB 11; Length 101;  
Best Local Similarity 100.0%; Pred. No. 2.5e-16;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 atggaactggaatgtgtataattacagaagaagaaggagagcttagtcagagaaga 172  
|||||  
Db 1 ATGGAAGTGAAGTGTATATATACAGAGAAACAGGAGAGCTTAGTGCAGAGGA 60  
|||||

QY 173 gacgaagtgcagagcggaacagcatccttagcttctcatat 213  
|||||  
Db 61 GAGCACTGTGAGAGGCGCAACAGCATCTTAGTCTTCATAT 101  
|||||

RESULT 9  
AC094509 80892 bp DNA linear HTG 20-DEC-2001  
LOCUS AC094509/c  
DEFINITION Rattus norvegicus clone CH230-4E17, \*\*\* SEQUENCING IN PROGRESS \*\*\*  
ACCESSION AC094509  
VERSION AC094509.3 GI:17941241  
KEYWORDS  
HTG: HTGS\_PHASE1.  
SOURCE Norway rat.  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 80892)  
REFERENCE  
1 (bases 1 to 80892)  
AUTHORS Muzny D.M., Adams C., Adio-Oduola B., Ali-osman F.R., Allen C.,  
Alsbrooks S.L., Amaratunga H.C., Are J.R., Banks T., Barbara J.,  
Benton J., Blinze K., Blankenburg K., Bonini D., Bouck J.,  
Bowles S., Blevins M., Brown E., Brown M., Bryant N.P., Butay C.,  
Burke P., Burkett C., Burrell K.L., Byrd N.C., Cartron T.F.,  
Carter M., Cavazos S.R., Chacko J., Chavez D., Chen G., Chen R.,  
Chen Z., Chowdhury I., Christopoulos C., Cleveland C.D., Cox C.,  
Coyle M.D., Dathorne S.R., David R., Davila M.L., Davis C.,  
Davy Carroll J., Dederich D.A., Delaney K.R., Delgado O.,  
Dem A.L., Ding Y., Dinh H.H., Douhwaite K.J., Draper H.,  
Dugan-Rocha S., Durbin K.J., Earnhart C., Edgar D., Edwards C.C.,  
Elhaj C., Escoto M., Falls T., Ferraguto D., Flagg N., Ford J.,  
Foster P., Frantz P., Gabisi A., Gao J., Garcia A., Garner T.,  
Garza N., Gill R., Gorrell J.H., Guevara M., Gunaratne P., Hale S.,  
Hamilton K., Harris C., Harris K., Hart M., Havlak P., Hawes A.,



Hernandez, J., Hernandez, O., Hodgson, A., Hognes, M., Holloway, C.,  
Hollins, B., Homsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J.,  
Jackson, L., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S.,  
Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvath, J.,  
Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L. C.,  
Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W.,  
Louliseed, H., Lozano, R. J., Lu, X., Lucier, A., Lucier, R., Luna, R.,  
Ma, J., Maheshwari, M., Mapa, P., Martin, R., Martindale, A.,  
Martinez, E., Massey, E., Mawhney, E., McLeod, M. P., Mendor, M.,  
Mei, G., Metzger, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K.,  
Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N.,  
Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S.,  
Ogih, M., Okwunonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B.,  
Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L. L.,  
Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojubokan, I., Rolfe, M.,  
Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoostari, N.,  
Sisson, I., Sodergren, E., Sonalke, T., Sparks, A., Stanley, H.,  
Stone, H., Sutton, A., Swatek, A., Tabor, P., Tamerisa, A., Tamerisa, K.,  
Tang, H., Tansey, J., Taylor, C., Taylor, T., Tellrod, B., Thomas, N.,  
Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R.,  
Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C.,  
Watlington, S., Williams, G., Williamson, A., Wleczek, R., Wooden, S.,  
Worley, K., Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D.,  
Weinstock, G., and Gibbs, R.

Direct Submission  
Unpublished  
2 (bases 1 to 80892)  
Worley, K. C.

Direct Submission  
Submitted (15-SEP-2001) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Dec 20, 2001 this sequence version replaced gi:15799320.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GAUD  
Center clone name: CH230-4E17  
----- Summary Statistics  
Assembly program: Phrap, version 0.990329first call to  
findphraplist  
Consensus quality: 56832 bases at least Q40  
Consensus quality: 63643 bases at least Q30  
Consensus quality: 67630 bases at least Q20  
Estimated insert size: 39832; sum-of-contigs estimation  
Quality coverage: 0x in Q20 bases; agarose-fp estimation  
Quality coverage: 0.4x in Q20 bases; sum-of-contigs estimation  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 50 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1  
2037 2036: contig of 2036 bp in length  
2137 2136: gap of unknown length  
5541 5540: contig of 3404 bp in length  
5641 5640: gap of unknown length  
7154 7154: contig of 1514 bp in length  
7255 7254: gap of unknown length  
10127 10126: contig of 2872 bp in length  
10227 10226: gap of unknown length  
12304 12304: contig of 1978 bp in length  
14032 14032: gap of unknown length  
14033 14132: gap of unknown length

14133 16701: contig of 2569 bp in length  
16702 16801: gap of unknown length  
18902 18931: contig of 2130 bp in length  
18932 19031: gap of unknown length  
19032 20670: contig of 1639 bp in length  
20671 20770: gap of unknown length  
20771 23399: contig of 2629 bp in length  
23400 23499: gap of unknown length  
23500 25649: contig of 2150 bp in length  
25650 25749: gap of unknown length  
25750 27707: contig of 1988 bp in length  
27708 27807: gap of unknown length  
27808 29187: contig of 1380 bp in length  
29188 29287: gap of unknown length  
29288 30599: contig of 1312 bp in length  
30600 30699: gap of unknown length  
30700 32012: contig of 1313 bp in length  
32013 32112: gap of unknown length  
32113 33114: contig of 1002 bp in length  
33115 33214: gap of unknown length  
33215 34609: contig of 1395 bp in length  
34610 34709: gap of unknown length  
34710 35952: contig of 1243 bp in length  
35953 36052: gap of unknown length  
36053 37791: contig of 1739 bp in length  
37792 37891: gap of unknown length  
37892 39196: contig of 1305 bp in length  
39197 39296: gap of unknown length  
39297 40805: contig of 1509 bp in length  
40806 40905: gap of unknown length  
40906 42389: gap of unknown length  
42389 42489: gap of unknown length  
42490 43940: contig of 1451 bp in length  
43941 44040: gap of unknown length  
44041 45314: contig of 1274 bp in length  
45315 45414: gap of unknown length  
45415 46689: contig of 1275 bp in length  
46690 46789: gap of unknown length  
46789 48037: contig of 1248 bp in length  
48038 48137: gap of unknown length  
48138 49753: contig of 1616 bp in length  
49754 49853: gap of unknown length  
49854 51676: contig of 1823 bp in length  
51677 51776: gap of unknown length  
51776 52788: contig of 1012 bp in length  
52789 52888: gap of unknown length  
52889 54070: contig of 1182 bp in length  
54071 54170: gap of unknown length  
54171 55190: contig of 1020 bp in length  
55191 55280: gap of unknown length  
55280 55391: gap of unknown length  
55391 56560: contig of 1270 bp in length  
56561 56660: gap of unknown length  
56661 58150: contig of 1490 bp in length  
58151 58250: gap of unknown length  
58251 59449: contig of 1199 bp in length  
59450 59549: gap of unknown length  
59550 61139: contig of 1530 bp in length  
61140 61239: gap of unknown length  
61240 62362: contig of 1123 bp in length  
62363 62462: gap of unknown length  
62463 63610: contig of 1148 bp in length  
63611 63710: gap of unknown length  
63711 64847: contig of 1137 bp in length  
64848 64947: gap of unknown length  
64948 66262: contig of 1315 bp in length  
66263 66362: gap of unknown length  
66363 67452: contig of 1090 bp in length  
67453 67552: gap of unknown length  
67553 68667: contig of 1315 bp in length  
68668 68967: gap of unknown length  
68968 70101: contig of 1134 bp in length  
70102 71230: gap of unknown length  
71230 71230: contig of 1029 bp in length

Query Match	11.9%	Score 65.8	DB 2	Length 80892
Best Local Similarity	66.7%	Pred. No. 2.6e-06		
Matches 108	Conservative 0	Mismatches 53	Indels 1	Gaps 1
QY 385	ctgcagagagcttcagcaagctglaaagaggccccgggctccgcagacagccagactagag -	443		
Db 11982	CTGCAGAGATTTTACAAATGTGACGTGACGAGTCTGTGTGGTGCAGAAAGCCAGATTA	CTGACG	11923	
QY 444	ncaaaagccagcttcctcagagctccagcggccgctgcgataccacgccttcctcgcaact	503		
Db 11922	ACAAAGCCCAATCTTCAAGGCTGATGCCCCCTCTCTGTGTCAGAGGACTTCTCGCAAC	ACTT	11863	
QY 504	catcatcatcagggccgctcatatgcatcagatgtggcag	545		
Db 11862	GATCATTCAGCAGGTGTTCCCTTCATGATTCACAGCCAGATTTTGGAG	11821		

LOCUS	RESULT 10
AK024793/c	
DEFINITION	2543 bp mRNA linear
ACCESSION	AK024793
VERSION	AK024793.1
KEYWORDS	GI:10437180
SOURCE	oligo capping; f1s (full insert sequence).
ORGANISM	Homo sapiens primary smooth muscle cells of human coronary artery
REFERENCE	cDNA to mRNA, clone_11b; CAS clone: CAS07548.
AUTHORS	Homo sapiens
TITLE	Eukaryota: Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;
JOURNAL	Mammalia: Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE	1 (bases)
AUTHORS	Watanabe, K., Kunagai, A., Itakura, S., Yamazaki, M., Tashiro, H.,
TITLE	Ota, T., Suzuki, Y., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T.,
JOURNAL	Nakamura, Y., Isogai, T. and Sugano, S.
REFERENCE	NEDO human cDNA sequencing project
AUTHORS	Unpublished (2000)
TITLE	2 (bases 1 to 2543)
JOURNAL	Sugano, S., Suzuki, Y., Ota, T., Obayashi, M., Nishi, T., Isogai, T.,
REFERENCE	Shibahara, T., Tanaka, T. and Nakamura, Y.
AUTHORS	Direct Submission
TITLE	Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
JOURNAL	University of Tokyo, Laboratory of Genome Structure Analysis, Human
REFERENCE	Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
AUTHORS	Japan (E-mail: cna@lims.u-tokyo.ac.jp, Tel: 81-3-5449-5286,
TITLE	Fax: 81-3-5449-5416)
JOURNAL	NEDO human cDNA sequencing project supported by Ministry of
REFERENCE	International Trade and Industry of Japan: cDNA full insert
AUTHORS	sequencing: Research Association for Biotechnology: cDNA library
TITLE	construction, 5'- & 3'-end one pass sequencing: Department of
JOURNAL	Virology and Human Genome Center, Institute of Medical Science,
REFERENCE	University of Tokyo (partly supported by Science and Technology

[illegible]

Query Match	Score	DB	Length	5241 bp	DNA	linear	PAT 28-SEP-2001
Best Local Similarity	65.2%	Pred. No. 7.8e-06;					
Matches 107; Conservative	0;	Mismatches 56;	Indels 1;	Gaps 1;			

Oy	444	-ncaagacgaatccctcgcagctccacgcccgcgttcgatcacacgcgccttcacgaact	502
Db	5173	CACAAAGCAGTCCCAAGGCTGCATTCACACTTCCCTGTGCCAAGGGAACTTCTCAGCAAAC	5114
Oy	503	tcatcatcatcaggccgcgttcatcgatccagttgtgcagc	546
Db	5113	TGATCATGAGCAGTGTTGCGCTTGATGCTCAGCACGATTWTGAGGC	5070

RESULT 12  
AC107883/c

LOCUS AC107883/c

DEFINITION Homo sapiens chromosome 15 clone CTD-2318C11 map 15, WORKING DRAFT SEQUENCE, 4 ordered pieces.

ACCESSION AC107883

VERSION AC107883.1 GI:18308607

KEYWORDS HTG; HTGS\_PHASE2; HTGS\_DRAFT; HTGS\_FULLTOP.

SOURCE human.  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 131197)  
Mammalia: Eutheria: Primates: Catarrhini: Homnidae: Homo.  
Britten, B., Linton, L., Nussbaum, C. and Lander, E.  
Homo sapiens chromosome 15, clone CTD-2318C11 unpublished  
2 (bases 1 to 131197)

AUTHORS Britten, B., Linton, L., Nussbaum, C., Lander, E., All, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.Y., Chazaro, B., Choebel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galaean, J., Gardyna, S., Glade, S., Gord, S., Guyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A., Karatas, A., Kells, C., Larocque, K., Lamazares, R., Landers, T., Lenoczky, J., Levine, R., Liu, G., MacLean, C., MacDonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Menues, L., Mihova, T., Molina, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, F., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnapack, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Palamas, J., Testafaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zemdek, L., Zimmer, A. and Zody, M.

JOURNAL Direct Submission  
Submitted (24-JAN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
All repeats were identified using RepeatMasker:  
Smilt, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RW/RepeatMasker.html>

COMMENT  
TITLE  
JOURNAL  
Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
Project Information  
Center project name: I22183  
Center clone name: 2318.C.11  
----- Summary Statistics -----  
Sequencing vector: Plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 130312 bases at least Q40  
Consensus quality: 130627 bases at least Q30  
Consensus quality: 130721 bases at least Q20  
Insert size: 127000; agarose-fp  
Quality coverage: 16.1 in Q20 bases; agarose-fp

```

Quality coverage: 15.6 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*   1 15983: contig of 15983 bp in length
*   * 15984 16083: gap of 100 bp
*   * 16084 46826: contig of 30743 bp in length
*   * 46827 46926: gap of 100 bp
*   * 90567 90567: contig of 43641 bp in length
*   * 90568 90667: gap of 100 bp
*   * 90668 131197: contig of 40530 bp in length.
FEATURES
source
    Location/Qualifiers
        1..131197
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="15"
            /map="15"
            /clone="CTD-2318C11"
            /clone_1ID="CTD Human BAC"
            /note="assembly_fragment"
                1..15983
            /note="assembly_fragment"
                16084..46826
            /note="assembly_fragment"
                46927..90567
            /note="assembly_fragment"
                90668..131197
            /note="assembly_fragment"
                90668..131197
BASE COUNT      34710 a 31848 c 31069 g 33269 t       301 others
ORIGIN
Query Match      11.4%; Score 63; DB 2; Length 131197;
Best Local Similarity 65.2%; Pred. NO. 1.6e-05;
Matches 107; Conservative 0; Mismatches 56; Indels 1; Gaps 1;
QY 384 gctgcaggagcgttcacgagcgctgaagaagcccccggcgctccgcagacgccagtactag 443
||||| |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
Db 55142 GCTGCAGAAATTTTTCACATAACACATGAGGGAGACTGTGTAGTGCAAAAAGCCAAATATCTGAG 55083
||||| |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
QY 444 -ncaaaagccagtcctccacagctccacagccccgcgcctcgcataccaccgcgccttccgnaact 502
||||| |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
Db 55082 CACAAGAACCAGTCTCAAGACGTGATTCCACCTTCCCTGCCAGGGAGCTTCTCAGCAACT 55023
||||| |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
QY 503 tcattatcatcaggcccgcttcatttgtgatccatttgccagc 546
||||| |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
Db 55022 TGATCATGAGCAGTGTTCGCTTGATGCTAAGCCAGTTTTGGAGC 54979
||||| |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |

RESULT 13
AC016693/c
LOCUS          AC016693              161492 bp     DNA             linear   HTG 07-JUL-2000
DEFINITION    Homo sapiens chromosome 15 clone RP11-94P14, WORKING DRAFT
SEQUENCE      17 unordered pieces.
ACCESSION     AC016693
VERSION       AC016693.4 GI:7230876
KEYWORDS      HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE        human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE     1 (bases 1 to 161492)
AUTHORS      Waterston,R.H.
TITLE        The sequence of Homo sapiens clone
JOURNAL       unpublished
REFERENCE     2 (bases 1 to 161492)
AUTHORS      Waterston,R.H.
```



```

REFERENCE
AUTHORS
2 (bases 1 to 164871)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
Barra,N., Bastien,V., Boguslavsky,L., Boukhgalter,B., Brown,A.,
Cammarata,A., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Faro,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J.,
Gardyna,S., Glude,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heatford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Karatas,A., LaRoque,K., Lamazares,R., Landers,T.,
Lehoczky,J., Levine,R., Liu,G., Maclean,C., Macdonald,P.,
Marquis,N., Mathews,C., McCarthy,M., McEwan,P., McKernan,K.,
McPheeters,R., Meldrum,J., Meneus,L., Mihova,T., Mienga,V.,
Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,
Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Roselti,M.,
Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
Sounez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Travers,M., Travis,N., Triggillo,J., Vassiliev,H., Viel,R., Vo,A.,
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.

TITLE
JOURNAL
COMMENT
Submitted (05-JAN-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Nov 29, 2001 this sequence version replaced g1:14192956.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center -----
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIRB
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information -----
Center project name: L11896
Center clone name: 94.P.14

----- Summary Statistics -----
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 163163 bases at least Q40
Consensus quality: 164030 bases at least Q30
Consensus quality: 164234 bases at least Q20
Insert size: 170000; agarose-IP
Quality coverage: 8.4 in Q20 bases; sum-of-coverage
Quality coverage: 8.7 in Q20 bases; sum-of-coverage

----- NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 55484: contig of 55484 bp in length
* 55485 55584: gap of 100 bp
* 55585 73029: contig of 17445 bp in length
* 73030 73129: gap of 100 bp
* 73130 103409: contig of 30280 bp in length
* 103410 103509: gap of 100 bp
* 103510 145248: contig of 41739 bp in length
* 145249 145348: gap of 100 bp
* 145349 164871: contig of 19523 bp in length.

FEATURES
Source
1..164871
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="15"
/map="15"
/clone="Rp11-94P14"

```

```

misc_feature
1..55484
/clone-lib="RpC1-11 Human Male BAC"
/notes="assembly-fragment
clone-end:SP6
vector-side:left"
misc_feature
55585..73029
/notes="assembly-fragment"
73130..103409
/notes="assembly-fragment"
103510..145248
/notes="assembly-fragment"
145349..164871
/notes="assembly-fragment
vector-end:T7
vector-side:right"
BASE COUNT 52135 a 31854 c 31794 g 48688 t 400 others
ORIGIN
Query Match 11.4% Score 63; DB 2; Length 164871;
Best Local Similarity 65.2% Pred. No. 1.7e-05;
Matches 107; Conservative 0; Mismatches 56; Indels 1; Gaps 1;

```

```

QY 384 gctgcagagagcttcagcagctgtaagaagggcccgcttcgcaagcgcaggtactgag 443
DB 8246 GCTGCAGAAATTTCACATACACTGAGGAGCTCTGACTGCGCAAAAGCCAAATCTGAG 8305
QY 444 -ncaagcagcttcagctcagctccagcccgcttcgacacccgcttcgcgcaact 502
DB 8306 CACAAAGCCAGCTCCTCAGAGCTGATTCACACTTCCTGTCAGGAGCACTTCTGCAAACT 8365
QY 503 tcatcatcatcagagcccgcttcagtgcagctgagctgagcagc 546
DB 8366 TGATCATGAGCAGCTGCTGCTGATGCTGATGCTGAGCCAGTTTGGAGC 8409

```

## RESULT 15

```

AC087477 189335 bp DNA linear HTG 11-FEB-2002
LOCUS Homo sapiens chromosome 15 clone Rp11-522B15 map 15, *** SEQUENCING
DEFINITION IN PROGRESS ***, 2 ordered pieces.
ACCESSION AC087477
VERSION AC087477.5 GI:16799026
KEYWORDS HTG; HTGS_PHASE2; HTGS_FULLTOP; HTGS_ACTIVERTIN.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

```

```

REFERENCE
1 (bases 1 to 189335)
AUTHORS Barra,N., Bastien,V., Boguslavsky,L., Boukhgalter,B., Brown,A.,
Cammarata,A., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Faro,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J.,
Gardyna,S., Glude,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heatford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Karatas,A., LaRoque,K., Lamazares,R., Landers,T.,
Lehoczky,J., Levine,R., Liu,G., Maclean,C., Macdonald,P.,
Marquis,N., Mathews,C., McCarthy,M., McEwan,P., McKernan,K.,
McPheeters,R., Meldrum,J., Meneus,L., Mihova,T., Mienga,V.,
Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,
Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Roselti,M.,
Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
Sounez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Travers,M., Travis,N., Triggillo,J., Vassiliev,H., Viel,R., Vo,A.,
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 189335)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
Barra,N., Bastien,V., Boguslavsky,L., Boukhgalter,B., Brown,A.,
Cammarata,A., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Faro,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J.,
Gardyna,S., Glude,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heatford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Karatas,A., LaRoque,K., Lamazares,R., Landers,T.,
Lehoczky,J., Levine,R., Liu,G., Maclean,C., Macdonald,P.,
Marquis,N., Mathews,C., McCarthy,M., McEwan,P., McKernan,K.,
McPheeters,R., Meldrum,J., Meneus,L., Mihova,T., Mienga,V.,
Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,
Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Roselti,M.,
Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
Sounez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Travers,M., Travis,N., Triggillo,J., Vassiliev,H., Viel,R., Vo,A.,
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.

```

